Discovery of new endogenous viral elements in campoplegine wasps (Ichneumonidae, Campopleginae, Campoplex) suggests a widespread nudivirus domestication event

Alexandra Cerqueira de Araujo, Matthieu Leobold, Marie Cariou, Bernardo Ferreira dos Santos, Rustem Uzbekov, Renato Ricciardi, Pierluigi Scaramozzino, Andrea Lucchi, Karine Musset, Jean-Michel Drezen, et al.

Discovery of new endogenous viral elements in campoplegine wasps (Ichneumonidae, Campopleginae, Campoplex) suggests a widespread nudivirus domestication event

Alexandra Cerqueira de Araujo[a], Matthieu Leobold[b], Marie Cariou[c], Bernardo Ferreira dos Santos[d], Rustem Uzbekov[ef], Renato Ricciardi[g], Pierluigi Scaramozzino[g], Andrea Lucchi[g], Karine Musset[b], Jean-Michel Drezen[b], Thibaut Josse[b], Elisabeth Huguet[b]

[a]INRAE, Oniris, BIOEPAR, 44300, Nantes, France, [b]Institut de Recherche sur la Biologie de l'Insecte (IRBI), UMR 7261, CNRS - Université de Tours, Tours, France, [c]Service Analyse de Données, Muséum National d'Histoire Naturelle, Centre National de la Recherche Scientifique, Paris, France, [d] Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Center for Integrative Biodiversity Discovery, Invalidenstraße 43, Berlin, 10115, Germany, [e]Université de Tours, Département des Microscopies, Tours, France, [f]Faculty of Bioengineering and Bioinformatics, Moscow State University, Moscow, Russia , [g]Department of Agriculture, Food and Environment (DAFE), University of Pisa, Italy.

Corresponding author: alexandra.cdea@gmail.com

Abstract

Viral endogenization is a widespread phenomenon that generally results in non-functional viral elements. In parasitoid wasps, nudiviruses have been involved in three independent domestication events and play a key role in parasitism success. In Campopleginae (Ichneumonidae), endogenized ichnovirus presence is widespread, with the exception of *Venturia canescens* (Ichneumonidae Campopleginae), which is known to harbour an endogenous nudivirus that enables the wasp to produce Virus-Like-Particles (VLPs). So far, in Campopleginae, the nudivirus endogenization event was an isolated case restricted to *V. canescens.* Using third generation sequencing technologies, microscopy and mass spectrometry, we described new cases of domesticated nudiviruses in campoplegine wasps of the genus *Campoplex*. We found endogenized viruses in *Campoplex* genomes which belong to the Alphanudivrius genus and derive from the same endogenization event as the virus found in *V. canescens,* suggesting that this integration event could be much more widespread in Campopleginae than previously thought. These nudiviruses are organised in highly conserved clusters, but this organisation appears to have changed relatively to free viruses during the domestication process. We showed that *Campoplex capitator* produces in its ovaries VLPs that are similar morphologically to *V. canescens* VLPs, with almost the same protein content, except for a strikingly different virulence proteins, a sign of different evolutionary paths taken to respond to distinct evolutive pressure. While ichnovirus remnants have been previously

found in *V. canescens*, no trace of ichnovirus could be detected in *Campoplex* genomes, meaning that the evolutionary history of viral integrations in Campopleginae is more complex than previously imagined.

Keywords: Nudivirus, Viral domestication, Campopleginae, Virus-Like-Particle, Parasitoid wasp, Ichnovirus

Introduction

The endogenization of viral sequences within eukaryotic genomes is a common phenomenon in the evolutionary history of eukaryotic organisms (Feschotte & Gilbert, 2012; Holmes, 2011). These viral sequences, known as endogenous viral elements (EVEs), involve all virus types likely to integrate into any eukaryotic genome (Ballinger et al., 2014; Bejarano et al., 1996; Belyi et al., 2010; Crochu et al., 2004; Delaroque et al., 1999; Feschotte & Gilbert, 2012; Irwin et al., 2021; Katzourakis & Gifford, 2010; H. Liu et al., 2010; Maori et al., 2007). A large proportion of EVEs are segmented and non-functional (Arbuckle et al., 2010; Katzourakis & Gifford, 2010; Tarlinton et al., 2006). There are, however, examples of viral integrations providing novel adaptive functions and selective advantage to the eukaryotic organism that carries them. For example, some EVEs can confer protection to their host against genetically related viruses (Armezzani et al., 2014; Frank & Feschotte, 2017; Horie & Tomonaga, 2019).

Other EVEs are also involved in placentation in mammals (Lavialle et al., 2013; Mangeney et al., 2007) and synaptic plasticity in animals (Ashley et al., 2018; Campillos et al., 2006).

In general, for viral endogenization events that provide hosts with novel functions, co-option typically involves individual viral genes rather than entire viral genomes. However, in some parasitoid wasp species, entire portions of large double-stranded DNA (dsDNA) viruses have been integrated and perform crucial functions to the wasp's life cycle. Indeed, the wasps use the virus-derived particles to deliver virulence proteins that suppress the immune response of the wasp's host and spark other physiological changes that ensure successful parasitization by the wasp larva (Bézier, Herbinière, et al., 2009; Burke et al., 2013; J.-M. Drezen et al., 2017; Strand & Burke, 2012, 2014; Volkoff et al., 2010).

At least seven viral endogenization events occurred in independent wasp lineages, each involving the domestication of large dsDNA viruses (Burke, 2019; Burke et al., 2021; Di Giovanni et al., 2020; J.-M. Drezen et al., 2022; Guinet et al., 2023). These viral endogenization events can be classified into different categories based on the nature of the virus that has been endogenized. The first one involves endogenized nudivirus sequences, leading to the production of bracoviruses (BVs) and Virus-Like-Particles (VLPs; particles made of viral envelopes containing wasp proteins and no DNA) (Bézier, Herbinière, et al., 2009; Burke et al.,

2018; Pichon et al., 2015). The second one involves the endogenization of a viral ancestor (characterised by a conserved set of genes) from a still uncharacterized virus family in two distinct Ichneumonidae lineages (Banchinae and Campopleginae), producing Ichnovirus (IV) particles (Béliveau et al., 2015; Gundersen-Rindal et al., 2013; Strand & Burke, 2014; Volkoff et al., 2010; Volkoff & Cusson, 2020). The production of these particles involves the expression of several intronless genes tightly clustered into wasp genome regions called IVSPERs (Ichnovirus Structural Protein Encoding Regions) (Lorenzi et al., 2019; Volkoff et al., 2010, 2012). The last viral endogenization category involves filamentous viruses related to hytrosaviruses, probably allowing wasps to produce VLPs (Burke et al., 2021; Di Giovanni et al., 2020; Guinet et al., 2023; Wey et al., 2020).

So far, three of the seven endogenization events described in parasitoid wasps correspond to the integration of nudivirus sequences. The Nudiviridae family, which diverged 220 MA from the well-known Baculoviridae family, comprises viruses infecting arthropods (Abd-Alla et al., 2008; Jehle et al., 2013; Thézé et al., 2011; Y. Wang et al., 2012). Integration of nudivirus sequences have been described in genome of arthropods including several insects (Bézier et al., 2015; Burand et al., 2012; R.-L. Cheng et al., 2014, 2020; S. Liu et al., 2020, 2021; Y. Wang et al., 2007; Y. Zhang et al., 2020), but, except in parasitoid wasps, these integrations were not

reported as virus domestication since the function of any of the viral sequences has not been studied.

The integration of sequences from a deltanudivirus (Petersen et al., 2022) in the ancestor of Braconidae Microgastrinae wasps 100 MYA led this lineage, currently comprising at least 46000 species, to produce BV particles, which are composed of a viral envelope enclosing capsids containing dsDNA circular molecules harbouring virulence genes (Bézier, Annaheim, et al., 2009; Herniou et al., 2013; Murphy et al., 2008; Thézé et al., 2011). More recently, the integration of alphanudivirus sequences was demonstrated in the braconid *Fopius arisanus* (Burke et al., 2018) and more unexpectedly in the ichneumonid *Venturia canescens* (Pichon et al., 2015). In these parasitoid wasps, the endogenization event led to the production of VLPs (Burke et al., 2018; Burke, 2019; Pichon et al., 2015). Several lines of evidence have suggested that the alphanudivirus integration event in *V. canescens* was more recent than the deltanudivirus integration event in an ancient braconid wasp (100 MYA) (Murphy et al., 2008; Pichon et al., 2015; Thézé et al., 2011). Indeed, first, the nudivirus integration event has been described up to now in a single campoplegine wasp species. Second, the viral genes integrated in *V. canescens* are less dispersed in the wasp genome than the BV genome in braconid wasps (Bézier, Annaheim, et al., 2009; Burke et al., 2014; Mao et al., 2023; Pichon et al., 2015), possibly because fewer genomic rearrangements had time to occur in Campopleginae. It has also been shown that in VcENV, losses of certain core nudivirus functions (*i.e.*: proteins

109 involved in capsid formation) are due to pseudogenization involving accumulation of
110 mutations rather than complete loss of genes or genomic regions (Leobold et al., 2018). The
111 fact that pseudogenized nudiviral sequences have been maintained in VcENV and are not
112 completely eroded could be an indication that the nudivirus integration is more recent in this
113 species (Leobold et al., 2018).

114

115 So far, *Venturia canescens* was the only described representative of Campopleginae wasps
116 with an integrated alphanudivirus (Burke, 2019; Burke et al., 2021; Pichon et al.,
117 2015). Because *V. canescens* belongs to the subfamily Campopleginae, known to harbour
118 ichnoviruses, initially the VLPs were thought to derive from IVSPER expression (Reineke et al.,
119 2006). However, sequencing of the wasp's genome identified only a few pseudogenized
120 IVSPER genes, and it was formally demonstrated that the VLPs are in fact of nudiviral origin
121 (Pichon et al., 2015). The alphanudivirus integrated in the Campopleginae wasp *V. canescens*,
122 named "Venturia canescens Endogenous Nudivirus" (VcENV), allows the production VLPs,
123 which consist of a membrane enveloping virulence proteins that protect the wasp eggs from
124 the lepidopteran host immune system after oviposition (Bedwin, 1979; Feddersen et al., 1986;
125 Pichon et al., 2015; Reineke et al., 2006; Rotheram, 1967, 1973; Salt, 1965). Three virulence
126 proteins have been described in VcVLPs: VLP1, VLP2 and VLP3, which are respectively a Rho
127 GTPase activating protein, a phospholipid hydroperoxide glutathione peroxidase and a

neprilysin (Asgari et al., 2002; Hellers et al., 1996; Pichon et al., 2015; Reineke et al., 2002; Theopold et al., 1994). Hence, whilst two genera of Nudivirus have been endogenized in different wasp lineages, the evolutionary trajectories they have undertaken since integration have given rise to strikingly different particles but with similar functions (Burke, 2019; J.-M. Drezen et al., 2017).

In Campopleginae, nudiviral integration is only known from *V. canescens*. Since whole genome data is available only for a few species, it is still unclear whether this event is really restricted to one or a few species or if it is more widespread in campoplegine wasps. To investigate this issue, we selected a genus, *Campoplex*, that is putatively closely related to *V. canescens* (Santos et al., 2022).

In this study, we obtained the complete genome sequence of two *Campoplex* species, one of which corresponds to *Campoplex capitator*, the main parasitoid of the vineyard pest *Lobesia botrana* (J. Moreau et al., 2010; Papura et al., 2016; Thiéry et al., 2011). These species and their endogenized nudiviruses offered the opportunity to compare domesticated virus genomes and the produced particles to better understand the early mechanisms at the basis of viral domestication and whether these domestication events followed the same evolutionary trajectories.

Results

Two new Campopleginae wasp genomes of the *Campoplex* genera

To investigate mechanisms involved in nudiviral domestication in Campopleginae wasps, we aimed to obtain genomes of *Campoplex* species, which are phylogenetically close to *V. canescens*. For this purpose, we sequenced the genome of *C. capitator* and *Campoplex nolae* using long-read and short-read sequencing technologies. The genome assembly generated for *C. capitator* is 261 Mb long and composed of 630 contigs (generated by PacBio sequencing), while that of *C. nolae* is 218 Mb long and composed of 4 593 contigs (generated by Illumina and Nanopore sequencing) (see detailed information in supplementary data Table S1). The quality of these genomes, with low levels of fragmentation (N50 = 7,8 Mb for *C. capitator* and 1,8 Mb for *C. nolae*; see Table S1) and a good genome completeness (96% and 94% of complete hymenopteran BUSCOs for *C. capitator* and *C. nolae* respectively; see Table S2) allowed the description of integrated viruses (see next paragraph). Gene models of *C. capitator* were predicted using transcriptomic data obtained after RNA sequencing of wasp samples. As no transcriptomic data was available for *C. nolae*, gene models were generated using protein sequences of *C. capitator* predicted previously. In total, 11 288 and 13 929 genes models have been predicted in *C. capitator* and *C. nolae* respectively.

The endogenized nudivirus genomes within the *Campoplex* genomes

We identified 47 genes of viral origin in both wasp genomes (Table1 and Table S3). Out of the 32 nudivirus core genes, 23 have been found in the *Campoplex* genomes (Table 1). Among these genes, some encode key viral functions in baculoviruses such as viral transcription (*p47* and *lef* genes), DNA amplification (*helicase*), envelope composition/infectiosity (*Ac81, p33* and *pif* genes) and virion morphogenesis (*vlf-1*). These 47 genes are grouped into 4 clusters in both genomes (Figure 1 for *C. capitator* and Figure S1 for *C. nolae*). These clusters are dense in viral genes and are identical in gene content and order between the two species (Figure1, Figure S1 and Table 2).

In *V. canescens*, remnants of some nudivirus genes, notably those involved in DNA amplification and capsid formation, have been found as pseudogenes (Leobold et al., 2018). Therefore, for each missing nudivirus gene, we searched for remnant sequences in both *Campoplex* genomes. Several pseudogenes have been identified, some located inside or around the virus clusters (Figure 1 and Figure S1), others such as *38K* located outside (Table 1). Genes that are typically involved in DNA amplification (*integrase, FEN-1* and *DNApol*) and capsid formation (*p6.9, 38K* and *vp39*) in baculoviruses have been found pseudogenized in both *Campoplex* species (Table 2), except for *38K* which couldn't be found in *C. nolae*, probably due to the higher fragmentation level of this genome.

No traces of Ichnovirus remnants in *Campoplex* genomes

As remnants of Ichnovirus were found in *V. canescens* (Pichon et al., 2015), we also searched for sequences of IVSPER genes in the *Campopex* genomes. In both *Campoplex* genomes, no trace of ichnovirus could be found using the IVSPER protein sequences (see material and method for details).

The endogenized nudiviruses in *Campoplex* correspond to Alphanudiviruses

Viral sequences found in both *Campoplex* genomes share high similarity with nudivirus sequences (Table S3). We therefore aimed to place these new endogenized viruses in the context of the phylogeny of Naldaviricetes, a class of virus comprising Baculoviridae, Nudiviridae, Hytrosaviridae and Nimaviridae. Using protein sequences from 34 genes (21 baculoviruses core genes and 13 nudivirus core genes, see Table S4 for more details), the phylogenetic analysis places these endogenized viruses inside the Nudivirus family and as sister to VcENV, which belongs to the Alphanudivirus genus (Figure 2). Due to their placement in the Nudivirus family, we called these genomes CcapiENV and CnolaENV for Campoplex capitator Endogenous Nudivirus and Campoplex nolae Endogenous Nudivirus respectively. These two new virus genomes are from now on the closest relatives of VcENV.

Evidence for a single nudivirus endogenization event in Campopleginae wasps

Given the phylogenetic position of these integrated alphanudiviruses in *Campoplex* (Figure 2), we then investigated whether these nudiviruses resulted from the same integration event as the one for VcENV. Wasp genes around the virus clusters were found to be the same in both *C. capitator* and *V. canescens* species, indicating that the endogenization event occurred at the same place in both genomes (Figure 3). This observation confirms that these viruses come from the same endogenization event.

Evolution of the endogenized nudiviruses in Campopleginae wasps

We compared the three endogenized nudivirus genomes in Campopleginae wasps to gain insight on the mechanisms that are operating during the early processes of virus domestication. The gene content in these viruses is almost identical, with only a few differences lying in the gene position, in the number of copies for certain genes and in the identified pseudogene (Table 2). However, while all other genes are present in all genomes, *vlf-1* was found only in CcapiENV and CnolaENV, this gene being found in a pseudogenized form in VcENV (Table 2). In baculoviruses, VLF-1 is a nucleocapsid protein that participates in virion maturation, in the integration and excision of the viral genome and in viral transcription (McLachlin & Miller, 1994; Vanarsdall et al., 2006; Yang & Miller, 1998, 1999). Duplicated

217 copies of *Ac81* and *OrNVorf47-like* have been found only in VcENV (Table 2), indicating that
218 those copies could have originated from duplication events occurring after speciation. Copies
219 of *pif-5* have been found in all genomes, some in similar locations inside the clusters (Table 2
220 and Figure 3), indicating that these copies might come from a duplication event that had
221 happened in the common ancestors of all these wasps. The organisation in clusters is also
222 highly preserved among these endogenized viruses in Campopleginae wasps, with CcapiENV
223 clusters 1 and 2 being almost identical to the VcENV clusters 3 and 2 (Figure 3). CcapiENV
224 cluster 4 and its VcENV equivalent (cluster 4) seems to have endured some genomic shuffling,
225 but not enough to erase syntenies between the two genomes (Figure 3).

226

227 Impact of endogenization on nudivirus gene organisation

228 Integrated alphanudivirus genomes were compared to free virus genomes in order to evaluate
229 the impact of viral domestication after integration. Gene order among free nudiviruses is fairly
230 well conserved (Figure 4 A), but gene order was globally lost following viral endogenization
231 (Figure 4 B), suggesting that loss of certain syntenic regions might be part of the viral
232 domestication process.

233 We also compared the endogenized alphanudiviruses in Campopleginae wasps to other
234 endogenized alphanudiviruses such as FaENV (Fopius arisanus Endogenous Nudivirus) and

NlENV (Nilaparvata lugens Endogenous Nudivirus). First, as described before, the viral genome structure in Campopleginae wasps is conserved after integration (Figure 4 C). However, for each independent alphanudivirus integration, the genomes seem to have been shuffled differently (Figure 4 C). Viral genome shuffling after integration might be part of a viral domestication process, as developed in Burke and collaborators (Burke et al., 2018; see discussion).

Selection pressures acting on nudivirus genes

To estimate the selection pressures acting on *C. capitator*, *C. nolae* and *V. canescens* genes, dN/dS ratios were calculated for 8 787 cellular genes and the 41 endogenized nudivirus genes. For the vast majority of genes (13152 out of 13351 values), the dN/dS ratio is lower than 1 (Figure 5). Nonetheless, dN/dS ratios obtained for nudivirus genes are higher than the ones of cellular genes, with an average dN/dS ratio of 0.17 for cellular genes and 0.73 for nudivirus genes. Focusing on nudivirus gene evolution according to their function, genes putatively involved in DNA amplification (*helicase*) and viral transcription (*lefs* and *p47*) have a lower dN/dS ratio than genes putatively involved in infectivity and envelope protein production (*pifs*, *vp91*, *p33* and *p74*) (Figure 5). Gauthier and colleagues had also found equivalent results on nudiviral genes associated with Braconid wasps (Gauthier et al., 2018, 2021). Furthermore,

duplicated nudiviral genes in the three Campopleginae species have a rather high dN/dS ratio, notably the copies of *OrNVorf41-like* and *pif-5* (Table S5).

CcapiENV gene expression

Virus gene expression activity within the wasps has been investigated through transcriptome analyses performed on *C. capitator* samples of pools of organs from male and female individuals. Because virus genes participate in the production of virus-derived particles in the ovaries for several wasp species, notably in *V.canescens* (J. Drezen et al., 2006; Pichon et al., 2015; Stoltz, 1990), we assumed that CcapiENV genes would be mostly expressed in *C. capitator* ovaries. As expected, all CcapiENV genes are expressed in adult wasp ovaries (Figure S2 and Table S6). Transcripts from some nudivirus genes are also visible in venom glands. The head-thorax samples of female wasps appear to be almost free of viral transcripts, with only minute amounts of transcripts being measured. CcapiENV genes are practically not expressed in males, neither in testis nor in head-thorax samples, although testis samples show a slightly higher number of transcripts than head-thorax samples (Table S6). The *pif-2* gene appears to be an exception as it is weakly expressed in head-thorax samples of both males and female wasps.

Regulatory sequences have also been studied in *C. capitator* in order to understand regulation mechanisms leading to the expression of endogenized nudivirus genes. However, no congruent results could be observed as the type of promoter (baculovirus early, baculovirus late, VcENV late) does not necessarily match the expected expression kinetics of the CcapiENV genes based on the transcription kinetic described for their VcENV orthologs (Cerqueira de Araujo et al., 2022; Table S7).

C. capitator produces VLPs that are similar in morphology to those produced by *V. canescens*

After identifying the endogenized nudivirus in the *Campoplex* genomes, we investigated whether *C. capitator* produces VLPs as described in *V. canescens*. Electron microscopy revealed the presence of particles in the calyx region of *C. capitator* ovaries, a specialised portion of the ovaries that produces virus-derived particles in braconid wasps and Ichneumonid wasps (Bézier, Annaheim, et al., 2009; Burke & Strand, 2012; Pichon et al., 2015; Volkoff et al., 2010). These particles in *C. capitator* are referred to as CcapiVLPs (Figure 6), and are almost morphologically identical to VcVLPs (Pichon et al., 2015). CcapiVLPs are composed of a membrane enclosing an electron-dense body (Figure 6E) and are massively present in the lumen (Figure 6 A and B) while only a few are visible in the cytoplasm and cell nuclei (Figure 6

C). Virogenic stromas, electron-dense masses composed of proteins that form viral particles in other wasps (Bézier, Herbinière, et al., 2009; Pichon et al., 2015), are visible in calyx cell nuclei (Figure 6C), along with VLPs and empty envelopes (Figure 6F). As described for VcVLPs (Pichon et al., 2015), CcapiVLPs appear to be be produced inside the nucleus by virogenic stromas and transit through the nuclear membrane and then through the cytoplasmic membrane by budding to finally end in the calyx lumen (Figure 6 D,E and F).

C. capitator and *V. canescens* VLPs have a similar composition but the virulence protein contents are different

We proceeded to determine whether the protein composition of the VLPs and their content were the same between *C. capitator* and *V. canescens*. For CcapiVLPs, 28 proteins of nudivirus origin were detected by mass spectrometry (Table S8). In the majority of cases, proteins contained in VcVLPs deriving from VcENV genes were also found in CcapiVLPs. For instance, all the VcVLP proteins involved in envelope formation in baculoviruses were found in purified CcapiVLPs. These include all the PIF proteins, P74, VP91 and P33 proteins. However, OrNvorf46-like and a copy of PIF-5 proteins are present in VcVLPs but not in CcapiVLPs (Table S8). Similarly, OrNVorf18-like, OrNVorf47-like, OrNVorf79-like, OrNVorf90-like and

OrNVorf120-like proteins appear to be present in CcapiVLPs but were not detected in VcVLPs (Table S8).

A large number of non-viral proteins, which should include virulence proteins, could be identified by mass spectrometry within the VLPs (Table S9). Proteins described in VcVLPs were first sought in the CcapiVLP protein dataset to determine whether the same proteins are enveloped in VLPs in both wasp species. The VLP1 (PHGPx), VLP2 (RhoGAP) and VLP3 (Neprilysin) proteins of VcVLPs were not found in CcapiVLPs. Candidate VLP virulence proteins were selected according to their peptide abundance (PSM for Peptide-Spectrum Matches) value and only proteins with a PSM value higher than 50 were retained (Table S10). Some proteins that have been described in parasitoid wasp venoms (Inwood et al., 2023; S. Moreau & Asgari, 2015; Poirié et al., 2014), such as a serpin, a peroxiredoxin, a calreticulin and an aminopeptidase, have been found in CcapiVLPs. Further investigation on the serpin showed that the gene encoding this protein is a member of a multigenic family, which comprises 15 members in *Campoplex spp*. and 4 in *V. canescens*. While the four copies found in *V. canescens* have orthologous sequences in *Campoplex spp.*, the serpin identified in CcapiVLPs is more distant, with only one ortholog found in *C. nolae* (Figure S3). This serpin appears to be specific to *Campoplex* species and therefore represents an interesting virulence protein candidate.

Discussion

The sequencing of two new genomes of *Campoplex* wasps has provided the opportunity to describe, together with the previous sequencing of the *V. canescens* wasp genome (Mao et al., 2023; Pichon et al., 2015), the early events involved in the domestication of alphanudiviruses that these wasps have endogenized. Our results convincingly show that the endogenized nudiviruses in the *Campoplex* species correspond to the same endogenization event as the one first described in *V. canescens*. The few and compact viral gene clusters, together with the presence of visible traces of pseudogenes, strongly suggest that this endogenization event is more recent than the one described in Microgastrinae braconid wasps. Furthermore, strong synteny is evident between the three endogenized alphanudiviruses, which contrasts with total lack of synteny with the known free nudiviruses for which sequences are available. This result suggests that a very early process required in viral domestication involves major genomic rearrangements, in accordance with similar observations made in the case of FaENV (Burke et al., 2018). Comparisons between the endogenized nudiviruses in Campopleginae wasps reveal not only conserved synteny of genes but also very high gene similarities, suggesting that evolutionary constraints may be acting to conserve gene functions, to enable coordinated expression leading to VLP formation. Finally, VLPs produced by *V. canescens* and *C. capitator* wasps are very similar, but genes coding for the VLP proteins involved in infectiosity are undergoing higher selection pressures and the VLP

virulence protein contents are totally different, illustrating that selection pressures exerted by wasp's host are most probably acting and resulting in tailored-virulence protein contents within VLPs.

A recent and unique endogenization event of an alphanudivirus in Campopleginae wasps

CcapiENV and CnolaENV, the nudiviruses integrated in *C. capitator* and *C. nolae* genomes respectively, are composed of 47 nudiviral genes each. These viruses belong to the Nudivirus family, especially to alphanudiviruses, and are the closest known relative of VcENV, the only endogenized nudivirus that had been previously described in Campopleginae wasps (Figure 2). Comparing the insertion location of CcapiENV and VcENV, the same wasp genes in the same order have been found around the viral clusters in both species (Figure 3). These observations indicate that these two new viruses come from the same endogenization event as the one of VcENV, which occurred in the common ancestor of the wasps harbouring them.

The integration of the alphanudivirus in Campopleginae appears to have occurred more recently than the endogenisation of BV in Braconidae. First, the wasp phylogeny indicates that the subfamily Campopleginae is a younger lineage than the microgastroid complex (Burke et al., 2021; Murphy et al., 2008; Sharanowski et al., 2021) and not all representatives of the

359 Campopleginae family appear to possess nudivirus (such as wasps of the genus *Hyposoter* or
360 *Dusona*) (Burke et al., 2021), indicating that nudivirus integration is not ancestral to the
361 Campopleginae and occurred later in a specific Campopleginae lineage. Furthermore, in both
362 *Campoplex* species, the distribution of the endogenized virus genes shows that the viral
363 genome is not very dispersed, forming four highly conserved clusters within the wasp
364 genomes (Figure 1). This organisation in clusters has been previously described in VcENV (Mao
365 et al., 2023; Pichon et al., 2015) and in FaENV (Burke et al., 2018), which have five and nine
366 clusters respectively. In bracoviruses, nudivirus clusters have also been characterised but virus
367 genes are more dispersed in the wasp genomes (Bézier, Annaheim, et al., 2009; Burke et al.,
368 2014, 2018; Gauthier et al., 2021). These observations would indicate that alphanudiviruses
369 integrated in parasitoid wasps have undergone fewer rearrangement events than the BV
370 integrated in the Braconidae. Also, it was possible to find pseudogenized nudiviral genes in
371 *Campoplex* wasps and in *V. canescens* (Leobold et al., 2018), whereas no pseudogenes were
372 characterised for BVs, which again suggests that the alphanudivirus integration event in
373 Campopleginae wasps is more recent because traces of gene loss are still detectable.

374

Genome rearrangement seems to be part of the viral domestication process

Genomic rearrangements appear to be a part of the viral domestication process in parasitoid wasps. Indeed, the viral organisation of nudivirus genes of the Alphanudivirus genus is very well conserved among free Alphanudiviruses viruses infecting arthropods (Figure 4 A). While this viral organisation is kept in free viruses, the genomes of endogenized nudiviruses in Campopleginae wasps appear to have lost this well conserved viral organisation, with their genome being rearranged in their host genome (Figure 4 B). This phenomenon of genomic rearrangement after integration is shared by the viruses that have been independently endogenized in different lineages of insects and produce integrated viruses with distinct organisations (Figure 4 C). A similar observation was made in FaENV where the viral genome has been shuffled after endogenization in the wasp genome (Burke et al., 2018). It has been suggested that these early rearrangement events occurring after integration could lead to the deactivation of the virus replication process by scattering the virus genome and could correspond to a founder effect in the process of domestication (Burke et al., 2018).

Key viral functions and the organisation in clusters were kept after domestication

CcapiENV, CnolaENV and VcENV share almost exactly the same content in genes (Table 2). The 22 core nudivirus genes found in VcENV were also identified in *Campoplex* endogenized

393 viruses (Table 1). Genes necessary for VLP production were kept in these Campopleginae
394 wasps. It has been shown for example that genes such as *lef* genes that are involved in viral
395 transcription are particularly well conserved among wasps that have endogenized a nudivirus
396 (Bézier, Annaheim, et al., 2009; Bézier, Herbinière, et al., 2009; Burke, 2019; Burke et al., 2018;
397 Burke & Strand, 2012; J.-M. Drezen et al., 2022; Pichon et al., 2015; Strand & Burke, 2012).
398 Silencing of these genes leads to the suppression or reduction of particle production in *V.*
399 *canescens* and *Microplitis demolitor* (Burke et al., 2013; Cerqueira de Araujo et al., 2022).
400 Furthermore, in Campopleginae, genes involved in viral transcription (*p47* and *lef* genes) and
401 DNA amplification (*helicase*) are under strong conservative selection pressures compared to
402 genes typically involved in envelope formation and infectiosity (*p33*, *Ac81* and *pif* genes)
403 (Figure 5). The same observation was made in braconid wasps harbouring a BV and in
404 *Leptopilina* wasps harbouring an endogenized filamentous virus (Di Giovanni et al., 2020;
405 Gauthier et al., 2021), showing that the viral transcription function is globally maintained in
406 all these viral domestication events. Genes coding for envelope components (such as *pif*
407 genes) have also been conserved in BV, FaENV and Campopleginae endogenized nudiviruses,
408 although their sequences differ more than the ones of *lef* genes for instance (Bézier,
409 Annaheim, et al., 2009; Bézier, Herbinière, et al., 2009; Burke, 2019; Burke et al., 2018; Burke
410 & Strand, 2012; Di Giovanni et al., 2020; J.-M. Drezen et al., 2022; Gauthier et al., 2021; Pichon
411 et al., 2015; Strand & Burke, 2012). In baculoviruses, PIFs proteins are involved in the entry of

412 viral particles into the cells of the lepidopteran host's digestive tract (Boogaard, 2018; Kikhno
413 et al., 2002; Zheng et al., 2017). As they form the interface between the particle and the host
414 cells, these proteins are therefore likely to diverge depending on the host that the particles
415 infect.

416 Regarding the viral genome structure, the endogenized virus genome structure is organised in
417 clusters of viral genes in Campopleginae wasps, in which the gene order was preserved after
418 wasp diversification. VcENV, CcapiENV and CnolaENV are all composed of three main
419 conserved nudiviral clusters, with observable syntenies between species (Figure 3, Figure 4).
420 For instance, the composition and gene order of the clusters are the same in both CcapiENV
421 and CnolaENV (Figure 1 and Figure S1). Comparing CcapiENV with VcENV, genes of CcapiENV
422 clusters 1 and 2 have exactly the same order as genes in VcENV clusters 3 and 2 (Figure 3).
423 Nonetheless, the order of viral genes between CcapiENV cluster 4 and VcENV clusters 4 and 5
424 is less conserved, indicating that certain recombination events took place after speciation of
425 the parasitoid wasps hosting them (Figure 3). In contrast, pseudogenes identified in
426 *Campoplex* and *V. canescens* genomes are dispersed compared to functional genes (Table S3)
427 (Leobold et al., 2018; Mao et al., 2023).

Loss of genes not essential for VLP production

Remnants of the majority of genes that have been pseudogenized in VcENV can be found in CcapiENV and CnolaENV. Some pseudogenes, such as *OrNVorf130-like*, were not found in CcapiENV or CnolaENV. As pseudogenes can be very degraded, it can sometimes be difficult to detect them, but it is also possible that these genes have been lost by deletion as described in bracoviruses (Bézier, Annaheim, et al., 2009; Burke et al., 2014).

Genes involved in capsid formation in baculoviruses have been lost for the most part (see *vlf-1* below) in all Campopleginae wasps (Table 2). More specifically, FaENV and BVs have kept genes such as *38K* and *vp39* (Bézier, Annaheim, et al., 2009; Burke et al., 2014, 2018), genes that are involved in baculoviruses in nucleocapsid assembly and DNA packaging (Danquah et al., 2012; Katsuma & Kokusho, 2017; Lai et al., 2018; Wu et al., 2006), but these genes are pseudogenized in *Campoplex* wasps and *V. canescens* (Leobold et al., 2018; present study). Genes involved in DNA packaging such as *integrase* can be found in BVs but were lost by pseudogenization in Campopleginae wasps producing VLPs, which are devoid of DNA (Burke, 2019; Burke et al., 2013, 2018; J.-M. Drezen et al., 2022; Pichon et al., 2015).

Specific adaptations of Campopleginae wasps to produce VLPs

During virus domestication, the alphanudivirus genome has been well conserved among Campopleginae wasps. Gene content and genome structure are similar between CcapiENV/CnolaENV and VcENV. However, notable differences can be observed between these genomes, probably being hallmarks of different adaptation to the wasp's hosts.

Vlf-1 is present in *Campoplex,* but not in *V. canescens*

One of the major differences between the two new genomes and VcENV is the presence of *vlf-1* in CcapiENV and CnolaENV, while this gene has been found pseudogenized in VcENV (Table 2). In baculoviruses, *vlf-1* is involved in several processes. The VLF-1 protein is an integrase that binds to specific DNA sites, catalyses DNA rearrangements and participates in the integration and excision of the viral genome (McLachlin & Miller, 1994). This protein, which is included in the nucleocapsid (Yang & Miller, 1998), appears to be necessary for virion maturation (Vanarsdall et al., 2006) and is involved in the transcription of genes expressed at very late stages of the viral infection (Yang & Miller, 1999). In CcapiENV, this gene does not appear to code for a structural VLP protein (Table 1 and Table S8), but is expressed specifically in wasp ovaries (Figure S2 and Table S6). The *vlf-1* gene may have a function in the transcription of late expressed genes and could possibly explain why we are not able to identify the same promoter sequences in CcapiENV and VcENV.

Duplication of nudivirus genes

In giant viruses—dsDNA viruses which exhibit a high level of gene duplications (Filée, 2018; Legendre et al., 2018) the expansion and reduction of the number of genes seems to be an adaptive process to respond to new constraints imposed by the host (Boyer et al., 2011; Elde et al., 2012). Among nudivirus duplicated genes in Campopleginae wasps (Table 2), *Ac81* and *OrNVorf47-like* have been found duplicated only in VcENV (Mao et al., 2023; Pichon et al., 2015). In addition, two extra copies of *pif-5* were found in VcENV and an extra copy of *OrNVorf41-like* was identified in CcapiENV and CnolaENV (Figure 1, Figure S1 and Figure 3). The *OrNVorf41-like* gene belongs to the 11K gene family, which comprises the *Ac145* and *Ac150* genes of the Autographa californica nucleopolyhedrovirus (AcMNPV). *Ac145* and *Ac150* appear to be involved in oral infectivity, although they are not essential for infection (J.-H. Zhang et al., 2004). The action of these genes is host-dependent: the deletion of *Ac145* reduces the infectivity of AcMNPV in the host *Trichoplusia ni,* but not in the host *Heliothis virescens*. The deletion of *A150* alone has no effect on either host. By deleting both *Ac145* and *Ac150*, the infectivity of AcMNPV was decreased in both *T. ni* and *H. virescens* (Lapointe et al., 2004). It has been proposed that *Ac145* is specifically involved in the initiation of primary infection while *Ac150* is more involved in the systemic spread of the infection (Beperet et al., 2015). The extra copy of *OrNVorf41-like* identified in CcapiENV and CnolaENV may therefore enable infection of different wasp's hosts compared to *V. canescens*.

CcapiVLPs and VcVLP do not share the same VLP content

Regarding the virulence proteins contained in VLPs, *V. canescens* produces VLPs containing three virulence proteins: VLP1, VLP2 and VLP3. In *C. capitator*, no VcVLP virulence protein counterparts, nor protein of similar functions, could be found inside CcapiVLPs. Wasp virulence proteins can be encoded from IV and BV circles or be produced in venom glands (Burke & Strand, 2014; J.-M. Drezen et al., 2017; Poirié et al., 2009). We therefore searched for virulence protein candidates which have similar functions to virulence proteins from other domesticated viruses or venom proteins. Of these virulence protein candidates, we found a serine protease inhibitor (serpin), which is a protein that has also been detected in venom of several parasitoid wasps and in MdBV (Microplitis demolitor Bracovirus) (Beck & Strand, 2007; Inwood et al., 2023; Moreau & Asgari, 2015; Poirié et al., 2014; Quicke & Butcher, 2021). These proteins have been shown to inhibit the proteolytic cascade leading to the production of phenoloxidase (PO), involved in the melanisation reaction (Beck & Strand, 2007; Colinet et al., 2009).

Other proteins described in wasp venoms have also been identified in the proteomic analysis of CcapiVLP components such as a peroxiredoxin, a calreticulin, and an aminopeptidase. Peroxiredoxin enzymes protect against oxidative stress and have also been identified in the venom of the parasitoid wasp *Anisopteromalus calandrae* (Perkin et al., 2015). In the context of parasitoid wasp-host interactions, this type of enzyme may protect the parasitoid eggs from

500 detrimental reactive oxygen species generated during oviposition and the encapsulation
501 response. Calreticulins have been identified in several parasitoid wasps (Cha et al., 2015; De
502 Graaf et al., 2010; Etebari et al., 2011; Fang et al., 2011; Perkin et al., 2015; L. Wang et al.,
503 2013; G. Zhang et al., 2006). Several studies have suggested calreticulin works as an
504 antagonist, competing for binding sites with host calreticulins, thus inhibiting hemocyte
505 encapsulation (Cha et al., 2015; L. Wang et al., 2013; G. Zhang et al., 2006). Aminopeptidase
506 activity or proteins have been detected in the venoms of several parasitoid wasps
507 (Becchimanzi et al., 2020; Dani et al., 2005; Inwood et al., 2023; Mathé-Hubert et al., 2016;
508 Teng et al., 2017), although the functions of these enzymes in the interaction are not clear
509 they could be involved in general degradation of host tissues.

510

511 It has often been described that genes encoding virulence proteins are part of multigenic
512 families expanded in species harbouring them. For instance, protein tyrosine phosphatases
513 (PTP) genes have been found largely duplicated in several BVs in which they code for virulence
514 proteins (Serbielle et al., 2012). In parasitoid wasps, some venom proteins are derived from
515 duplicated genes (Casewell et al., 2013; Colinet et al., 2014; Fry et al., 2009; Wong & Belov,
516 2012), conferring a selective advantage linked to the diversification of venom proteins. Among
517 all possible candidates (Table S10), we investigated whether these proteins are encoded from
518 genes included in expanded gene families in *Campoplex* species, with no obvious counterparts

519 in *V. canescens* genome. In CcapiVLPs, the serpin is the only protein that meets the criteria
520 described above (see results), establishing it as a major candidate. Thus, it would be
521 interesting to investigate its function using RNA interference to validate its virulence protein
522 function.

523

524 ## No support for ichnovirus replacement hypothesis in *Campoplex* species

525 In the study of Pichon et al. (2015), sequences that appeared to correspond to degraded
526 remnants of the ichnovirus replication machinery (IVSPERs) were detected in *V. canescens*
527 genome; this result, combined with a preliminary phylogeny inferred from 28S rRNA
528 sequences, led the authors to hypothesise that *V. canescens* may have undergone a process
529 of "endogenous virus replacement", in which an ancestrally present ichnovirus was lost
530 following the acquisition of the endogenous alphanudivirus. In the present study, we were
531 unable to corroborate these results (Pichon et al., 2015) as we could not find remnants of a
532 past ichnovirus integration in the *Campoplex* genomes. Similarly, Burke and collaborators
533 (Burke et al., 2021) did not find IVSPER sequences in a genome of *Dusona sp.*, one of the
534 closest relative to *Campoplex* species (Sharanowski et al., 2021). Previous results and our
535 results suggest that the ichnovirus integration might have happened in the Hyposoter -

Campoletis clade, while the Campoplex - Dusona clade have endogenized an alphanudivirus, which had been partially lost in *Dusona sp.* (see Sharanowski et al. (2021) for phylogeny).

To conclude, using third generation sequencing technology, it has been possible to describe new endogenous viruses in parasitoid wasps. The new genomes of two Campopleginae wasps harbour an alphanudivirus which derive from the same endogenization event as the one of VcENV in the common wasp ancestor. The description of these new endogenized nudiviruses sheds light on early evolutionary processes of viral domestication. Comparison with the closely related *V. canescens* species shows that there are minute differences between the genomes of endogenized nudiviruses in Campopleginae wasps. However, major differences can be seen, such as the virulence protein content of VLP which appears to be strikingly different between *C. capitator* and *V. canescens*, showing early adaptation to wasp's hosts in the context of viral domestication.

Material and methods

Parasitoid model

Campoplex capitator wasps were obtained after emergence from the lepidopteran host *Lobesia botrana*, a vineyard pest that can grow on *Daphne gnidium* (Loni et al., 2016). Caterpillars of *L. botrana* were sampled from *D. gnidium* shrubs in the natural reserve of Migliarino-San Rossore-Massaciuccoli (Tuscany, Italy). Samplings were carried out in early July of 2019 and 2021, with wasps emerging between June and August (Scaramozzino et al., 2018). *L. botrana* nests were placed individually in test tubes and each tube was monitored every day. Emerging *C. capitator* were used for experiments.

To produce enough wasps for our experiments, a *C. capitator* rearing was maintained as described in (Benelli et al., 2020; Lucchi et al., 2018): *L. botrana* larvae were kept in plastic boxes (20cm D x 15cm W x 10cm H) with a nutritious substrate (30g of agar, 60g of sugar, 50g of alfalfa, 36g of Brewer yeast, 25g of Wesson salt, 180g of wheat sprout, 80g of casein, 4g of sorbic acide, 15g of Wanderzahnt vitamins, 2.5g of cholesterol, 2.5g of tetracyclin, 5mL of propionic acid (99,5%), 2mL of linoleic acid (95%), 5mL of olive oil and 1.5mL of distilled water). After the emergence, adults were kept in the same box for reproduction where the substrate was removed and droplets of honey were added. Once eggs were visible on the box surface, adults were removed and nutritious substrate was added to feed the future larvae. *Campoplex*

568 *capitator* females and males were meanwhile kept in plexiglass chambers (40cm D x 25cm W
569 x 30cm H) with honey to allow their reproduction. L2 to L4 host larvae were put in the
570 chambers for two hour parasitism periods and then individually reared into small boxes (5cm
571 Ø x 5cm H) with nutritious substrate. These boxes were monitored every day to check for wasp
572 emergence and honey was added when larvae were about to pupate. All insects were kept at
573 25°C and 45% of humidity, with a 16:8 day/night photoperiod.

574 Meanwhile, a female specimen of *Campoplex nolae* was collected in the wild by active
575 sweeping at Powdermill Nature Reserve, Rector county, Pennsylvania, USA, and kept in liquid
576 nitrogen until the time of DNA isolation.

577

578 DNA extraction and sequencing

579 For *C. capitator*, DNA was extracted from one haploid male at emergence using the
580 MagAttract® HMW DNA Kit (Qiagen). The male wasp tissues were disrupted in a 2 mL tube
581 containing 200 µL of 1X DPBS. 20 µL of proteinase K, 4 µL of RNAse A and 150µL of buffer AL
582 were added to the tube. The tube was then mixed carefully and incubated at 56°C at 900 rpm
583 for 2 hours. The manufacturer's protocol was followed from step 4 to step 16. Total DNA was
584 finally eluted in 100 µL of TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) and quantified using
585 the Qubit™ DNA HS assay kit (Invitrogen) and the Qubit2.0 fluorometer (Invitrogen). DNA

586 quality and contamination were checked by measurement of 260/280 and 260/230 OD ratios
587 using the Varian Cary® 50 Scan spectrophotometer. Total DNA was then stored at 4°C. The
588 library was prepared and sequenced using the PacBio Sequel II technology by the Gentyane
589 platform (France).

590 For *C. nolae*, extraction was performed using the MagAttract® HMW DNA Kit (Qiagen),
591 following the manufacturer's protocol with the whole specimen immersed in extraction buffer
592 and proteinase K overnight. An additional purification step using Agencourt AMPure XP beads
593 (Beckman Coulter, Brea, U.S.A.) at 0.5X was used to remove small fragments. Genomic
594 libraries were prepared for parallel sequencing in short-read (Illumina) and long-read
595 (Nanopore) platforms: the Illumina library was prepared using the Kapa HyperPrep kit (Kapa
596 Biosystems, Wilmington, MA). A-tailing, end-repair and ligation reactions were performed at
597 a quarter volume relative to the standard protocol, while library amplification was performed
598 at full volume. Custom, dual-indexing adapter-primers were used to allow for in silico de-
599 multiplexing of each sample (Glenn et al., 2016). Following stub ligation and PCR, adapter-
600 dimers were removed by performing a 0.8X bead cleaning using AMPure beads. The library
601 was pooled with other samples at equimolar concentrations and sequenced at 4 nM as single
602 lanes on Illumina NovaSeq 6000 S4 platform (2x150; Illumina Inc., San Diego, CA). The library
603 for long-read data was prepared using the Nanopore ligation kit, complemented by reagents
604 from the NEBNext Ultra II library prep kit (New England Biolabs, Ipswich, Massachusetts,

605 U.S.A.), and then sequenced at a MinION platform (Oxford Nanopore Technologies, Oxford,
606 U.K.) using the manufacturer's specifications.

607 For both wasp species, raw data are available in NCBI (accession number of
608 the BioProject: PRJNA936130).

609

610 RNA extraction and sequencing

611 Emerging *C. capitator* wasps were dissected in order to obtain ovaries where CcapiVLP are
612 produced. Venom glands and head-thorax of female adults and testis and head-thorax of male
613 adults were also dissected in order to investigate expression of CcapiENV genes outside of the
614 calyx region and in males. For all samples, two pools were analysed. Pools were composed of
615 5 individuals for head-thorax for both males and females, 21 female venom glands, 5 ovary
616 pairs and 15 male testis pairs. Total RNA was extracted from the 10 pools in total
617 using TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. Only the
618 isopropanol step was replaced by cold absolute ethanol (3 volumes per sample) and total RNA
619 was finally resuspended in 30µL of RNAse free water. RNA quantity of each pool was
620 quantified using the Qubit™ RNA HS assay kit (Invitrogen) designed for the Qubit2.0
621 fluorometer (Invitrogen). Samples of extracted RNA were then stored at -80°C.

For genome annotation, total RNA was extracted from two pools of four entire emerging female wasps using the NucleoSpin®RNA kit (Macherey-Nagel). Total RNA was finally resuspended in 40µL of RNAse free water. RNA quantity of the two pools was quantified using the Qubit™ RNA HS assay kit (Invitrogen) designed for the Qubit2.0 fluorometer. Samples of extracted RNA were then stored at -80°C.

The sequencing of library preparations and the generation of paired-end reads were performed on an Illumina platform by the Novogene company (UK). Raw data are available in NCBI (accession number of the BioProject: PRJNA936130).

Genome assembly and annotation

For *C. capitator*, genome and K-mer sizes were estimated beforehand using KmerGenie (Chikhi & Medvedev, 2014). Raw reads were then corrected and trimmed using respectively Canu–correct and Canu–trim modules (Canu software v1.8) (Koren et al., 2017) in order to improve the base accuracy and to retrieve high quality sequence portions only. The genome was then assembled using the Canu–assemble module. The newly assembled genome was polished using Racon (v1.4.21) with the longest 80X raw reads to further improve the base accuracy using the longest raw reads. In parallel, *C. nolae* raw reads from Nanopore sequencing were filtered using Filtlong (v.0.2.1) before being assembled using Flye (v.2.9-b1768) (Kolmogorov

et al., 2020). The produced assembly was then polished with Hypo (v.1.0.3.flye) using the Illumina reads. Scripts for *C. nolae* genome assembly are available at https://gitlab.in2p3.fr/marie.cariou/campoplex_assembly/.

The assembled genome completeness was assessed with BUSCO (Benchmarking Universal Single-Copy Orthologs, v5.4.2) (Manni et al., 2021) using Insecta and Hymenoptera lineage datasets (odb10) and genome contamination was checked with Blobtools (v1.1.1) (Laetsch & Blaxter, 2017) for both genomes.

For both genomes, repeated elements were annotated *de novo* using the TEdenovo pipeline comprised in the REPET pipeline (v2.5) (Flutre et al., 2011). Two rounds of the TEannot pipeline (contained in the REPER pipeline) were then used to annotate repeated elements using existing databases. Repeated elements were then masked using Bedtools (v.2.29.1) (Quinlan & Hall, 2010) before gene prediction. Gene models were generated for *C. capitator* using BRAKER-2 (v2.1.6) (Brůna et al., 2021) with assembled transcriptomes by spades (v3.15.5) (Bushmanova et al., 2019). Models were then functionally annotated using InterProScan (v5.53-87.0) (Blum et al., 2021) and blastp (NCBIblast+ v2.13.0) with the nr database (downloaded the 14th of April 2023), before being polished by AGAT (v1.0.0) (Dainat et al., 2023). Proteins were extracted from gene models using the agat_sp_extract_sequences.pl tool from AGAT software. Finally, the genome completeness was assessed again using the protein sequences with BUSCO.

In total, 44 CcapiENV genes were annotated with the automated annotation process. Using the gene models predicted previously, additional viral genes were then annotated by reciprocal blasts: translated sequences from predicted genes were aligned using blast tools on the NCBI nr database filtered for sequences of endogenous and exogenous nudiviruses. Aligned sequences were then extracted and realigned on the complete NCBI nr database. The final Blast outputs were filtered to keep only sequences which aligned on sequences of endogenous and exogenous nudiviruses reported in the database.

To annotate nudivirus genes on the *C. nolae* genome, open reading frames were predicted using orfipy (v0.0.4) (Singh & Wurtele, 2021). Nudivirus genes were then annotated by reciprocal blasts using translated sequences from predicted ORFs. Miniprot (v0.9-0) (H. Li, 2023) was also used to align *C. capitator* proteins on *C. nolae* genome and validate previously annotated nudivirus genes. Miniprot gene models were also used for gene clustering in the dN/dS ratio analysis.

Pseudogenized genes were identified by reciprocal blast: tblastn was used first to map nudivirus proteins on the wasp genomes. Hits were then verified by blastx to map potential nudiviral nucleotide sequences on the NCBI nr database. Proteins corresponding to Nucleotide sequences mapping back on nudivirus were kept and incomplete and non-transcribed sequences were considered as pseudogenes.

Ichnovirus sequences were searched using a home-made pipeline where genomic IVSPER sequences of *Hyposoter dydimator* and *Glypta fumiferanae* were first aligned on *Campoplex* genome portions (the genomes were segmented into chunks of 500,000 bases). Aligning chunks were then aligned on the NCBI nr database. These hits were filtered to keep only hits aligning on IVSPER proteins that are absent from parasitoid wasps which do not bear an Ichnovirus. These "IVSPER" hits were then aligned back on the nr database and only the first hit per query was retained.

CO1 and nudivirus gene alignments

Because the endogenized nudivirus genomes were highly similar in both *Campoplex* species, we collected evidences to confirm that the *Campoplex* wasps sequenced for this study are part of different species. To determine if the genomes of *C. capitator* and *C. nolae* come from distinct species, whole genome alignments were performed using D-Genies (v1.4.0) (Cabanettes & Klopp, 2018). A CO1 (cytochrome oxydase 1) sequence alignment was also performed in order to measure divergence between the two genomes. CO1 orthologs were searched for in *C. capitator* and *C. nolae* genomes by aligning ORFs on CO1 sequence of *Campoplex* species reported in BOLD database (Ratnasingham & Hebert, 2007). Candidates were then aligned together (*C. capitator* candidates vs *C. nolae* candidates) before being

aligned on the NCBI database. Endogenized nudivirus were also aligned by blastp in order to assess the protein divergence between the two species. All results for this part are available in the supplementary materials.

Endogenized nudivirus evolution

To perform the virus phylogeny, protein sequences of one nimavirus, three hytrosaviruses, sixty-nine baculoviruses and twenty nudiviruses (including CcapiENV and *C. nolae* sequences) were first aligned together in order to find orthologous genes. Results were then manually corrected using the knowledge found in the literature (Bézier, Annaheim, et al., 2009; Bézier et al., 2015; Bézier, Herbinière, et al., 2009; Burand et al., 2012; Burke, 2019; J.-M. Drezen et al., 2022). Protein sequences of orthologous genes were then aligned using mafft (v7.487) (Katoh et al., 2019) with the following parameters: maxiterate=1000 (maximum number of iterations), genafpair (Altschul algorithm (Altschul, 1998)), ep=0 (offset value). Alignments were then trimmed by TrimAL (v1.4.rev15) (Capella-Gutiérrez et al., 2009) with the gap threshold parameter fixed at 0.6. For each gene, a tree was finally generated with IQ-TREE 2 (v2.1.4) (Minh et al., 2020) with the following parameters: m=TEST (use the best model found), B=1000 (number of bootstrap iterations), alrt=1000 (number of SH-aLRT replicates). Protein sequences of genes that were not supporting the accepted phylogeny (at the family level)

were removed from the further analysis. Finally, 21 baculovirus core genes and 13 alphanudivirus core genes were kept for the virus tree construction. Protein sequences of these genes were aligned and trimmed as described previously. Sequences were then concatenated using FASconCAT-G (v1.0) (Kück & Longo, 2014) and a tree was constructed with iqtree2 with the same parameters as previously described. The tree image was then built using ape (v5.7-1), treeio (v1.24.0) (L.-G. Wang et al., 2020) and ggtree (v3.8.0) (Yu et al., 2017) packages with R (v4.2.1).

A whole genome synteny analysis was performed between *V. canescens*, *C. capitator* and *C. nolae* using D-Genies. The cluster synteny figure (Figure 3) between *C. capitator* and *V. canescens* was built using R and parity plots (Figure 4) showing syntenies among nudiviruses were generated with Excel (Microsoft v16).

To estimate the evolutionary constraint occurring on groups of genes according to their function, protein sequences of *C. capitator*, *V. canescens* and *C. nolae* were first clustered using OrthoMCL (v2.0.9) (L. Li et al., 2003) and scripts written by Arun Seetharam (https://github.com/ISUgenomics/common_scripts). Clustered protein sequences were then aligned using ClustalOmega (v1.2.4) and aligned sequences were converted into codon alignement using PAL2NAL (v14.1) (Suyama et al., 2006). Phylip trees were generated using IQ-TREE 2 with the following parameters: -m JTT (Jones-Taylor-Thornton model (Jones et al., 1992)). Codeml (Paml v4.10.6) was finally used to calculate dN and dS values.

Regulatory sequences in CcapiENV

In the same fashion as described in Cerqueira de Araujo et al. (2022), promoter sequences regulating the early and late expression of viral genes in baculoviruses (Chen et al., 2013; Passarelli & Guarino, 2007; Rohrmann, 2014) were sought in the 300 bp upstream of CcapiENV genes. The "early" promoters described in Baculoviruses include a TATA box (TATA[A/T]T[A/T]) and a sequence motif (CA[G/T]T or CGTCG) placed among the 40 nucleotides after the TATA box. The late promoters include the baculovirus motif ([A/T/G]TAAG) and the Heliothis zea nudivirus 1 motif (TTATAGTAT) (C.-H. Cheng et al., 2002). We also considered the "late" motif ([G/T][A/T][A/G]A[A/T]ATAG[T/A]) described in Alexandra Cerqueira 2022 for which 2 mismatches were allowed in the sequence search.

CcapiENV gene expression

Paired raw reads from ovaries, head-thorax, testes and venom gland of *C. capitator* wasps were first trimmed using Trimmomatic (v 0.38) (Bolger et al., 2014) and TruSeq3-PE adaptors. A quality check was then performed on trimmed reads using fastqc (v 0.11.9) before mapping. Reads were mapped on the genome using Hisat-2 (v2.2.1-3n) (Kim et al., 2019) after genome indexation. Mapped reads were then sorted using samtools (v1.16). Sorted reads were converted into counts by FeatureCounts from the Rsubread package (v2.14.1) (Liao et al.,

2019). Using the edgeR package (v 3.42.2) (Robinson et al., 2010), genes with a count being equal to less than 15 per sample were filtered from the dataset in order to remove non-expressed genes. . The dataset was then normalised by converting the read counts into TPM (transcripts per million). A Spearman correlation heatmap was performed to verify duplicate consistency (Figure S4).

CcapiVLP purification and protein sequencing

In order to purify Virus-Like-Particles, 61 adult female wasps of *C. capitator* were dissected 2 days after emergence. In total, 122 calyces were obtained and pooled in 1X DPBS. Calyces were then disrupted thanks to a pellet pestle and CcapiVLPs were purified as described in Pichon et al. (2015): disrupted calyces were centrifuged at 770 g, at 4°C for 10 minutes in order to separate CcapiVLPs and calyx fluid (supernatant) from cellular waste (pellet). The supernatant was then retrieved and transferred into a new tube. The new tube was then centrifuged at 15 400 g, at 4°C for 10 minutes. The supernatant was discarded and the pellet containing CcapiVLPs was suspended in 20 µL of 1X DPBS and stored at -20°C.

CcapiVLP proteins were then separated by SDS-polyacrylamide gel electrophoresis (12.5% polyacrylamide gel) under denaturing conditions. Separated bands were incubated for 1h with a Coomassie blue solution (10% acetic acid, 50% ethanol, 0.1% R250 Coomassie blue, q.s. to

768 1L of milliQ water) and 9 slices were cut from the gel. Slices were chopped into 1mm^3 cubes
769 before being stored at -20°C and analysed by the platform PIXANIM (France) using a
770 nanoUPLC_LTQ-VElos Pro Orbitrap Mass Spectrometer.

771 For each protein identified by the mass spectrometry, sites and domains were searched for
772 using InterProScan and gene ontology terms were attributed using sequence homology with
773 curated proteins from the Uniprot SWISSprot database (downloaded the 18th November
774 2022). Proteins with a peptide abundance higher than 50 were then retained as VLP
775 candidates.

776

777 Electronic microscopy of the calyx region

778 The calyx region of adult *C. capitator* wasp ovaries was observed under electronic microscopy.
779 Wasp ovaries were dissected two days after emergence and fixed in PrOx/EPON epoxy resin.
780 Samples were fixed, cut and observed with the same protocol as one used in Cerqueira de
781 Araujo and collaborators (Cerqueira de Araujo et al., 2022). The protocol is available in
782 Supplementary materials.

783

Acknowledgements

The authors would like to thank Cosci Francesca, Filippo Di Giovanni and Augusto Loni from the Department of Agriculture, Food and Environment (DAFE) of the University of Pisa (Italy) for the help provided during field wasp sampling. The authors would like to thank also Amine Chebbi who helped the first author on the transposable element annotation pipeline. We would like to thank the IBiSA Electron Microscopy Facility of the University of Tours and the University Hospital of Tours for their assistance. Research funding was provided by an ADS Research Grant from the Smithsonian Institution to BFS, who was also supported by a GGI Peter Buck postdoctoral fellowship during early stages of this work.

Data availability

Raw reads and assemblies are deposited on the NCBI database under BioProject PRJNA936130.

Reference

Abd-Alla, A. M. M., Cousserans, F., Parker, A. G., Jehle, J. A., Parker, N. J., Vlak, J. M., Robinson, A. S., & Bergoin, M. (2008). Genome Analysis of a *Glossina pallidipes* Salivary Gland Hypertrophy Virus Reveals a Novel, Large, Double Stranded Circular DNA Virus. *Journal of Virology*, *82*(9), Article 9. https://doi.org/10.1128/JVI.02588-07

Altschul, S. F. (1998). Generalized affine gap costs for protein sequence alignment. *Proteins: Structure, Function, and Genetics*, *32*(1), 88–96. https://doi.org/10.1002/(SICI)1097-0134(19980701)32:1<88::AID-PROT10>3.0.CO;2-J

Arbuckle, J. H., Medveczky, M. M., Luka, J., Hadley, S. H., Luegmayr, A., Ablashi, D., Lund, T. C., Tolar, J., De Meirleir, K., Montoya, J. G., Komaroff, A. L., Ambros, P. F., & Medveczky, P. G. (2010). The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proceedings of the National Academy of Sciences*, *107*(12), Article 12. https://doi.org/10.1073/pnas.0913586107

Armezzani, A., Varela, M., Spencer, T., Palmarini, M., & Arnaud, F. (2014). "Ménage à Trois": The Evolutionary Interplay between JSRV, enJSRVs and Domestic Sheep. *Viruses*, *6*(12), Article 12. https://doi.org/10.3390/v6124926

Asgari, S., Reineke, A., Beck, M., & Schmidt, O. (2002). Isolation and characterization of a neprilysin-like protein from *Venturia canescens* virus-like particles. *Insect Molecular Biology*, *11*(5), Article 5.

Ashley, J., Cordy, B., Lucia, D., Fradkin, L. G., Budnik, V., & Thomson, T. (2018). Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell*, *172*(1–2), Article 1–2. https://doi.org/10.1016/j.cell.2017.12.022

Ballinger, M. J., Bruenn, J. A., Hay, J., Czechowski, D., & Taylor, D. J. (2014). Discovery and Evolution of Bunyavirids in Arctic Phantom Midges and Ancient Bunyavirid-Like Sequences in Insect Genomes. *Journal of Virology*, *88*(16), Article 16. https://doi.org/10.1128/JVI.00531-14

Becchimanzi, A., Avolio, M., Bostan, H., Colantuono, C., Cozzolino, F., Mancini, D., Chiusano, M. L., Pucci, P., Caccia, S., & Pennacchio, F. (2020). Venomics of the ectoparasitoid wasp Bracon nigricans. *BMC Genomics*, *21*(1), 34. https://doi.org/10.1186/s12864-019-6396-4

Beck, M. H., & Strand, M. R. (2007). A novel polydnavirus protein inhibits the insect prophenoloxidase activation pathway. *Proceedings of the National Academy of Sciences*, *104*(49), Article 49. https://doi.org/10.1073/pnas.0708056104

Bedwin, O. (1979). An insect glycoprotein: A study of the particles responsible for the resistance of a parasitoid's egg to the defence reactions of its insect host. *Proceedings of the Royal Society of*

London. Series B. Biological Sciences, *205*(1159), Article 1159. https://doi.org/10.1098/rspb.1979.0065

Bejarano, E. R., Khashoggi, A., Witty, M., & Lichtenstein, C. (1996). Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. *Proceedings of the National Academy of Sciences*, *93*(2), Article 2. https://doi.org/10.1073/pnas.93.2.759

Béliveau, C., Cohen, A., Stewart, D., Periquet, G., Djoumad, A., Kuhn, L., Stoltz, D., Boyle, B., Volkoff, A.-N., Herniou, E. A., Drezen, J.-M., & Cusson, M. (2015). Genomic and Proteomic Analyses Indicate that Banchine and Campoplegine Polydnaviruses Have Similar, if Not Identical, Viral Ancestors. *Journal of Virology*, *89*(17), Article 17. https://doi.org/10.1128/JVI.01001-15

Belyi, V. A., Levine, A. J., & Skalka, A. M. (2010). Sequences from Ancestral Single-Stranded DNA Viruses in Vertebrate Genomes: The *Parvoviridae* and *Circoviridae* Are More than 40 to 50 Million Years Old. *Journal of Virology*, *84*(23), Article 23. https://doi.org/10.1128/JVI.01789-10

Benelli, G., Ricciardi, R., Romano, D., Cosci, F., Stefanini, C., & Lucchi, A. (2020). Wing-fanning frequency as a releaser boosting male mating success—High-speed video analysis of courtship behavior in *Campoplex capitator* , a parasitoid of *Lobesia botrana*. *Insect Science*, *27*(6), Article 6. https://doi.org/10.1111/1744-7917.12740

Beperet, I., Simón, O., Williams, T., López-Ferber, M., & Caballero, P. (2015). The "11K" gene family members sf68, sf95 and sf138 modulate transmissibility and insecticidal properties of Spodoptera frugiperda multiple nucleopolyhedrovirus. *Journal of Invertebrate Pathology*, *127*, 101–109. https://doi.org/10.1016/j.jip.2015.03.008

Bézier, A., Annaheim, M., Herbinière, J., Wetterwald, C., Gyapay, G., Bernard-Samain, S., Wincker, P., Roditi, I., Heller, M., Belghazi, M., Pfister-Wilhem, R., Periquet, G., Dupuy, C., Huguet, E., Volkoff, A.-N., Lanzrein, B., & Drezen, J.-M. (2009). Polydnaviruses of Braconid Wasps Derive from an Ancestral Nudivirus. *Science*, *323*(5916), Article 5916. https://doi.org/10.1126/science.1166788

Bézier, A., Herbinière, J., Lanzrein, B., & Drezen, J.-M. (2009). Polydnavirus hidden face: The genes producing virus particles of parasitic wasps. *Journal of Invertebrate Pathology*, *101*(3), Article 3. https://doi.org/10.1016/j.jip.2009.04.006

Bézier, A., Thézé, J., Gavory, F., Gaillard, J., Poulain, J., Drezen, J.-M., & Herniou, E. A. (2015). The Genome of the Nucleopolyhedrosis-Causing Virus from *Tipula oleracea* Sheds New Light on the Nudiviridae Family. *Journal of Virology*, *89*(6), Article 6. https://doi.org/10.1128/JVI.02884-14

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., … Finn, R. D. (2021). The InterPro protein

865 families and domains database: 20 years on. *Nucleic Acids Research*, *49*(D1), D344–D354.
866 https://doi.org/10.1093/nar/gkaa977
867 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence
868 data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170
869 Boogaard, B. (2018). An Advanced View on Baculovirus *per Os* Infectivity Factors. *Insects*, *9*(84), Article
870 84. https://doi.org/10.3390/insects9030084
871 Boyer, M., Azza, S., Barrassi, L., Klose, T., Campocasso, A., Pagnier, I., Fournous, G., Borg, A., Robert, C.,
872 Zhang, X., Desnues, C., Henrissat, B., Rossmann, M. G., La Scola, B., & Raoult, D. (2011).
873 Mimivirus shows dramatic genome reduction after intraamoebal culture. *Proceedings of the*
874 *National Academy of Sciences*, *108*(25), Article 25. https://doi.org/10.1073/pnas.1101118108
875 Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2021). BRAKER2: Automatic
876 eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein
877 database. *NAR Genomics and Bioinformatics*, *3*(1), Article 1.
878 https://doi.org/10.1093/nargab/lqaa108
879 Burand, J. P., Kim, W., Afonso, C. L., Tulman, E. R., Kutish, G. F., Lu, Z., & Rock, D. L. (2012). Analysis of
880 the Genome of the Sexually Transmitted Insect Virus *Helicoverpa zea* Nudivirus 2. *Viruses*, *4*(1),
881 Article 1. https://doi.org/10.3390/v4010028
882 Burke, G. R. (2019). Common themes in three independently derived endogenous nudivirus elements
883 in parasitoid wasps. *Current Opinion in Insect Science*, *32*, 28–35.
884 https://doi.org/10.1016/j.cois.2018.10.005
885 Burke, G. R., Hines, H. M., & Sharanowski, B. J. (2021). The Presence of Ancient Core Genes Reveals
886 Endogenization from Diverse Viral Ancestors in Parasitoid Wasps. *Genome Biology and*
887 *Evolution*, *13*(7), Article 7. https://doi.org/doi:10.1093/gbe/evab105
888 Burke, G. R., Simmonds, T. J., Sharanowski, B. J., & Geib, S. M. (2018). Rapid Viral Symbiogenesis via
889 Changes in Parasitoid Wasp Genome Architecture. *Molecular Biology and Evolution*, *35*(10),
890 Article 10. https://doi.org/10.1093/molbev/msy148
891 Burke, G. R., & Strand, M. R. (2012). Deep Sequencing Identifies Viral and Wasp Genes with Potential
892 Roles in Replication of *Microplitis demolitor* Bracovirus. *Journal of Virology*, *86*(6), Article 6.
893 https://doi.org/10.1128/JVI.06434-11
894 Burke, G. R., & Strand, M. R. (2014). Systematic analysis of a wasp parasitism arsenal. *Molecular*
895 *Ecology*, *23*(4), Article 4. https://doi.org/10.1111/mec.12648
896 Burke, G. R., Thomas, S. A., Eum, J. H., & Strand, M. R. (2013). Mutualistic Polydnaviruses Share
897 Essential Replication Gene Functions with Pathogenic Ancestors. *PLoS Pathogens*, *9*(5), Article
898 5. https://doi.org/10.1371/journal.ppat.1003348
899 Burke, G. R., Walden, K. K. O., Whitfield, J. B., Robertson, H. M., & Strand, M. R. (2014). Widespread

Genome Reorganization of an Obligate Virus Mutualist. *PLoS Genetics*, *10*(9), Article 9. https://doi.org/10.1371/journal.pgen.1004660

Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: A de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, *8*(9), giz100. https://doi.org/10.1093/gigascience/giz100

Cabanettes, F., & Klopp, C. (2018). D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, *6*, e4958. https://doi.org/10.7717/peerj.4958

Campillos, M., Doerks, T., Shah, P., & Bork, P. (2006). Computational characterization of multiple Gag-like human proteins. *Trends in Genetics*, *22*(11), Article 11. https://doi.org/10.1016/j.tig.2006.09.006

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Casewell, N. R., Wüster, W., Vonk, F. J., Harrison, R. A., & Fry, B. G. (2013). Complex cocktails: The evolutionary novelty of venoms. *Trends in Ecology & Evolution*, *28*(4), Article 4. https://doi.org/10.1016/j.tree.2012.10.020

Cerqueira de Araujo, A., Leobold, M., Bézier, A., Musset, K., Uzbekov, R., Volkoff, A.-N., Drezen, J.-M., Huguet, E., & Josse, T. (2022). Conserved Viral Transcription Plays a Key Role in Virus-Like Particle Production of the Parasitoid Wasp *Venturia canescens*. *Journal of Virology*, *96*(13), Article 13. https://doi.org/10.1128/jvi.00524-22

Cha, W. H., Kim, Y., & Lee, D.-W. (2015). Calreticulin in Cotesia plutellae suppresses immune response of Plutella xylostella (L.). *Journal of Asia-Pacific Entomology*, *18*(1), 27–31. https://doi.org/10.1016/j.aspen.2014.11.001

Chen, Y.-R., Zhong, S., Fei, Z., Hashimoto, Y., Xiang, J. Z., Zhang, S., & Blissard, G. W. (2013). The Transcriptome of the Baculovirus *Autographa californica* Multiple Nucleopolyhedrovirus in *Trichoplusia ni* Cells. *Journal of Virology*, *87*(11), Article 11. https://doi.org/10.1128/JVI.00194-13

Cheng, C.-H., Liu, S.-M., Chow, T.-Y., Hsiao, Y.-Y., Wang, D.-P., Huang, J.-J., & Chen, H.-H. (2002). Analysis of the Complete Genome Sequence of the Hz-1 Virus Suggests that It Is Related to Members of the *Baculoviridae*. *Journal of Virology*, *76*(18), Article 18. https://doi.org/10.1128/JVI.76.18.9024-9034.2002

Cheng, R.-L., Li, X.-F., & Zhang, C.-X. (2020). Nudivirus Remnants in the Genomes of Arthropods. *Genome Biology and Evolution*, *12*(5), Article 5. https://doi.org/10.1093/gbe/evaa074

Cheng, R.-L., Xi, Y., Lou, Y.-H., Wang, Z., Xu, J.-Y., Xu, H.-J., & Zhang, C.-X. (2014). Brown Planthopper Nudivirus DNA Integrated in Its Host Genome. *Journal of Virology*, *88*(10), Article 10.

https://doi.org/10.1128/JVI.03166-13

Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, *30*(1), Article 1. https://doi.org/10.1093/bioinformatics/btt310

Colinet, D., Anselme, C., Deleury, E., Mancini, D., Poulain, J., Azéma-Dossat, C., Belghazi, M., Tares, S., Pennacchio, F., Poirié, M., & Gatti, J.-L. (2014). Identification of the main venom protein components of *Aphidius ervi*, a parasitoid wasp of the aphid model *Acyrthosiphon pisum*. *BMC Genomics*, *15*(1), Article 1. https://doi.org/10.1186/1471-2164-15-342

Colinet, D., Dubuffet, A., Cazes, D., Moreau, S., Drezen, J.-M., & Poirié, M. (2009). A serpin from the parasitoid wasp Leptopilina boulardi targets the Drosophila phenoloxidase cascade. *Developmental & Comparative Immunology*, *33*(5), Article 5. https://doi.org/10.1016/j.dci.2008.11.013

Crochu, S., Cook, S., Attoui, H., Charrel, R. N., De Chesse, R., Belhouchet, M., Lemasson, J.-J., de Micco, P., & de Lamballerie, X. (2004). Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. Mosquitoes. *Journal of General Virology*, *85*(7), Article 7. https://doi.org/10.1099/vir.0.79850-0

Dainat, J., Hereñú, D., Dr. K. D. Murray, Davis, E., Crouch, K., LucileSol, Agostinho, N., Pascal-Git, Zollman, Z., & Tayyrov. (2023). *NBISweden/AGAT: AGAT-v1.2.0* (v1.2.0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.8178877

Dani, M. P., Edwards, J. P., & Richards, E. H. (2005). Hydrolase activity in the venom of the pupal endoparasitic wasp, Pimpla hypochondriaca. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *141*(3), 373–381. https://doi.org/10.1016/j.cbpc.2005.04.010

Danquah, J. O., Botchway, S., Jeshtadi, A., & King, L. A. (2012). Direct Interaction of Baculovirus Capsid Proteins VP39 and EXON0 with Kinesin-1 in Insect Cells Determined by Fluorescence Resonance Energy Transfer-Fluorescence Lifetime Imaging Microscopy. *Journal of Virology*, *86*(2), Article 2. https://doi.org/10.1128/JVI.06109-11

De Graaf, D. C., Aerts, M., Brunain, M., Desjardins, C. A., Jacobs, F. J., Werren, J. H., & Devreese, B. (2010). Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. *Insect Molecular Biology*, *19*(s1), 11–26. https://doi.org/10.1111/j.1365-2583.2009.00914.x

Delaroque, N., Maier, I., Knippers, R., & Mv°ller, D. G. (1999). Persistent virus integration into the genome of its algal host, Ectocarpus siliculosus (Phaeophyceae). *Journal of General Virology*, *80*(6), Article 6. https://doi.org/10.1099/0022-1317-80-6-1367

Di Giovanni, D., Lepetit, D., Guinet, B., Bennetot, B., Boulesteix, M., Couté, Y., Bouchez, O., Ravallec, M., & Varaldi, J. (2020). A Behavior-Manipulating Virus Relative as a Source of Adaptive Genes

970 for *Drosophila* Parasitoids. *Molecular Biology and Evolution*, *37*(10), Article 10.
971 https://doi.org/10.1093/molbev/msaa030
972 Drezen, J., Bézier, A., Lesobre, J., Huguet, E., Cattolico, L., Periquet, G., & Dupuy, C. (2006). The few
973 virus-like genes of *Cotesia congregata* bracovirus. *Archives of Insect Biochemistry and*
974 *Physiology*, *61*(3), 110–122. https://doi.org/10.1002/arch.20108
975 Drezen, J.-M., Bézier, A., Burke, G. R., & Strand, M. R. (2022). Bracoviruses, ichnoviruses, and virus-like
976 particles from parasitoid wasps retain many features of their virus ancestors. *Current Opinion*
977 *in Insect Science*, *49*, 93–100. https://doi.org/10.1016/j.cois.2021.12.003
978 Drezen, J.-M., Leobold, M., Bézier, A., Huguet, E., Volkoff, A.-N., & Herniou, E. A. (2017). Endogenous
979 viruses of parasitic wasps: Variations on a common theme. *Current Opinion in Virology*, *25*, 41–
980 48. https://doi.org/10.1016/j.coviro.2017.07.002
981 Elde, N. C., Child, S. J., Eickbush, M. T., Kitzman, J. O., Rogers, K. S., Shendure, J., Geballe, A. P., & Malik,
982 H. S. (2012). Poxviruses Deploy Genomic Accordions to Adapt Rapidly against Host Antiviral
983 Defenses. *Cell*, *150*(4), Article 4. https://doi.org/10.1016/j.cell.2012.05.049
984 Etebari, K., Palfreyman, R. W., Schlipalius, D., Nielsen, L. K., Glatz, R. V., & Asgari, S. (2011). Deep
985 sequencing-based transcriptome analysis of Plutella xylostella larvae parasitized by Diadegma
986 semiclausum. *BMC Genomics*, *12*(1), 446. https://doi.org/10.1186/1471-2164-12-446
987 Fang, Q., Wang, F., Gatehouse, J. A., Gatehouse, A. M. R., Chen, X., Hu, C., & Ye, G. (2011). Venom of
988 Parasitoid, Pteromalus puparum, Suppresses Host, Pieris rapae, Immune Promotion by
989 Decreasing Host C-Type Lectin Gene Expression. *PLoS ONE*, *6*(10), e26888.
990 https://doi.org/10.1371/journal.pone.0026888
991 Feddersen, I., Sander, K., & Schmidt, O. (1986). Virus-like particles with host protein-like antigenic
992 determinants protect an insect parasitoid from encapsulation. *Experientia*, *42*(11–12), Article
993 11–12. https://doi.org/10.1007/BF01946422
994 Feschotte, C., & Gilbert, C. (2012). Endogenous viruses: Insights into viral evolution and impact on host
995 biology. *Nature Reviews Genetics*, *13*(4), Article 4. https://doi.org/10.1038/nrg3199
996 Filée, J. (2018). Giant viruses and their mobile genetic elements: The molecular symbiosis hypothesis.
997 *Current Opinion in Virology*, *33*, 81–88. https://doi.org/10.1016/j.coviro.2018.07.013
998 Flutre, T., Duprat, E., Feuillet, C., & Quesneville, H. (2011). Considering Transposable Element
999 Diversification in De Novo Annotation Approaches. *PLoS ONE*, *6*(1), Article 1.
1000 https://doi.org/10.1371/journal.pone.0016526
1001 Frank, J. A., & Feschotte, C. (2017). Co-option of endogenous viral sequences for host cell function.
1002 *Current Opinion in Virology*, *25*, 81–89. https://doi.org/10.1016/j.coviro.2017.07.021
1003 Fry, B. G., Roelants, K., Champagne, D. E., Scheib, H., Tyndall, J. D. A., King, G. F., Nevalainen, T. J.,
1004 Norman, J. A., Lewis, R. J., Norton, R. S., Renjifo, C., & de la Vega, R. C. R. (2009). The

Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms. *Annual Review of Genomics and Human Genetics*, *10*(1), Article 1. https://doi.org/10.1146/annurev.genom.9.081307.164356

Gauthier, J., Boulain, H., van Vugt, J. J. F. A., Baudry, L., Persyn, E., Aury, J.-M., Noel, B., Bretaudeau, A., Legeai, F., Warris, S., Chebbi, M. A., Dubreuil, G., Duvic, B., Kremer, N., Gayral, P., Musset, K., Josse, T., Bigot, D., Bressac, C., … Drezen, J.-M. (2021). Chromosomal scale assembly of parasitic wasp genome reveals symbiotic virus colonization. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-020-01623-8

Gauthier, J., Gayral, P., Le Ru, B. P., Jancek, S., Dupas, S., Kaiser, L., Gyapay, G., & Herniou, E. A. (2018). Genetic footprints of adaptive divergence in the bracovirus of *Cotesia sesamiae* identified by targeted resequencing. *Molecular Ecology*, *27*(8), Article 8. https://doi.org/10.1111/mec.14574

Guinet, B., Lepetit, D., Charlat, S., Buhl, P. N., Notton, D. G., Cruaud, A., Rasplus, J.-Y., Stigenberg, J., De Vienne, D. M., Boussau, B., & Varaldi, J. (2023). Endoparasitoid lifestyle promotes endogenization and domestication of dsDNA viruses. *eLife*, *12*, e85993. https://doi.org/10.7554/eLife.85993

Gundersen-Rindal, D., Dupuy, C., Huguet, E., & Drezen, J.-M. (2013). Parasitoid polydnaviruses: Evolution, pathology and applications: Dedicated to the memory of Nancy E. Beckage. *Biocontrol Science and Technology*, *23*(1), Article 1. https://doi.org/10.1080/09583157.2012.731497

Hellers, M., Beck, M., Theopoid, U., Kamei, M., & Schmidt, O. (1996). Multiple alleles encoding a virus-like particle protein in the ichneumonid endoparasitoid *Venturia canescens*. *Insect Molecular Biology*, *5*(4), Article 4. https://doi.org/10.1111/j.1365-2583.1996.tb00098.x

Herniou, E. A., Huguet, E., Thézé, J., Bézier, A., Periquet, G., & Drezen, J.-M. (2013). When parasitic wasps hijacked viruses: Genomic and functional evolution of polydnaviruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1626), Article 1626. https://doi.org/10.1098/rstb.2013.0051

Holmes, E. C. (2011). The Evolution of Endogenous Viral Elements. *Cell Host & Microbe*, *10*(4), Article 4. https://doi.org/10.1016/j.chom.2011.09.002

Horie, M., & Tomonaga, K. (2019). Paleovirology of bornaviruses: What can be learned from molecular fossils of bornaviruses. *Virus Research*, *262*, 2–9. https://doi.org/10.1016/j.virusres.2018.04.006

Inwood, S. N., Harrop, T. W. R., & Dearden, P. K. (2023). The venom composition and parthenogenesis mechanism of the parasitoid wasp Microctonus hyperodae, a declining biocontrol agent. *Insect Biochemistry and Molecular Biology*, *153*, 103897.

https://doi.org/10.1016/j.ibmb.2022.103897

Irwin, N. A. T., Pittis, A. A., Richards, T. A., & Keeling, P. J. (2021). Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. *Nature Microbiology*, *7*(2), 327–336. https://doi.org/10.1038/s41564-021-01026-3

Jehle, J. A., Abd-Alla, A. M. M., & Wang, Y. (2013). Phylogeny and evolution of *Hytrosaviridae*. *Journal of Invertebrate Pathology*, *112*, S62–S67. https://doi.org/10.1016/j.jip.2012.07.015

Jones, D. T., Taylor, W. R., & Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*, *8*(3), 275–282. https://doi.org/10.1093/bioinformatics/8.3.275

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), Article 4. https://doi.org/10.1093/bib/bbx108

Katsuma, S., & Kokusho, R. (2017). A Conserved Glycine Residue Is Required for Proper Functioning of a Baculovirus VP39 Protein. *Journal of Virology*, *91*(6), Article 6. https://doi.org/10.1128/JVI.02253-16

Katzourakis, A., & Gifford, R. J. (2010). Endogenous Viral Elements in Animal Genomes. *PLoS Genetics*, *6*(11), Article 11. https://doi.org/10.1371/journal.pgen.1001191

Kikhno, I., Gutiérrez, S., Croizier, L., Croizier, G., & Ferber, M. L. (2002). Characterization of pif, a gene required for the per os infectivity of Spodoptera littoralis nucleopolyhedrovirus. *Journal of General Virology*, *83*(12), Article 12. https://doi.org/10.1099/0022-1317-83-12-3013

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), Article 8. https://doi.org/10.1038/s41587-019-0201-4

Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T. P. L., & Pevzner, P. A. (2020). metaFlye: Scalable long-read metagenome assembly using repeat graphs. *Nature Methods*, *17*(11), 1103–1110. https://doi.org/10.1038/s41592-020-00971-x

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive *k* -mer weighting and repeat separation. *Genome Research*, *27*(5), Article 5. https://doi.org/10.1101/gr.215087.116

Kück, P., & Longo, G. C. (2014). FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology*, *11*(1), Article 1. https://doi.org/10.1186/s12983-014-0081-x

Laetsch, D. R., & Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research*, *6*, 1287. https://doi.org/10.12688/f1000research.12232.1

1075 Lai, Q., Wu, W., Li, A., Wang, W., Yuan, M., & Yang, K. (2018). The 38K-Mediated Specific
1076 Dephosphorylation of the Viral Core Protein P6.9 Plays an Important Role in the Nucleocapsid
1077 Assembly of Autographa californica Multiple Nucleopolyhedrovirus. *Journal of Virology*, *92*(9),
1078 Article 9. https://doi.org/10.1128/JVI.01989-17
1079 Lapointe, R., Popham, H. J. R., Straschil, U., Goulding, D., O'Reilly, D. R., & Olszewski, J. A. (2004).
1080 Characterization of Two *Autographa californica* Nucleopolyhedrovirus Proteins, Ac145 and
1081 Ac150, Which Affect Oral Infectivity in a Host-Dependent Manner. *Journal of Virology*, *78*(12),
1082 Article 12. https://doi.org/10.1128/JVI.78.12.6439-6448.2004
1083 Lavialle, C., Cornelis, G., Dupressoir, A., Esnault, C., Heidmann, O., Vernochet, C., & Heidmann, T.
1084 (2013). Paleovirology of *'syncytins'*, retroviral env genes exapted for a role in placentation.
1085 *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1626), Article 1626.
1086 https://doi.org/10.1098/rstb.2012.0507
1087 Legendre, M., Fabre, E., Poirot, O., Jeudy, S., Lartigue, A., Alempic, J.-M., Beucher, L., Philippe, N.,
1088 Bertaux, L., Christo-Foroux, E., Labadie, K., Couté, Y., Abergel, C., & Claverie, J.-M. (2018).
1089 Diversity and evolution of the emerging Pandoraviridae family. *Nature Communications*, *9*(1),
1090 Article 1. https://doi.org/10.1038/s41467-018-04698-4
1091 Leobold, M., Bézier, A., Pichon, A., Herniou, E. A., Volkoff, A.-N., & Drezen, J.-M. (2018). The
1092 Domestication of a Large DNA Virus by the Wasp *Venturia canescens* Involves Targeted
1093 Genome Reduction through Pseudogenization. *Genome Biology and Evolution*, *10*(7), Article
1094 7. https://doi.org/10.1093/gbe/evy127
1095 Li, H. (2023). Protein-to-genome alignment with miniprot. *Bioinformatics*, *39*(1), btad014.
1096 https://doi.org/10.1093/bioinformatics/btad014
1097 Li, L., Stoeckert, C. J., & Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic
1098 genomes. *Genome Research*, *13*(9), 2178–2189. https://doi.org/10.1101/gr.1224503
1099 Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better
1100 for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*(8),
1101 Article 8. https://doi.org/10.1093/nar/gkz114
1102 Liu, H., Fu, Y., Jiang, D., Li, G., Xie, J., Cheng, J., Peng, Y., Ghabrial, S. A., & Yi, X. (2010). Widespread
1103 Horizontal Gene Transfer from Double-Stranded RNA Viruses to Eukaryotic Nuclear Genomes.
1104 *Journal of Virology*, *84*(22), Article 22. https://doi.org/10.1128/JVI.00955-10
1105 Liu, S., Coates, B. S., & Bonning, B. C. (2020). Endogenous viral elements integrated into the genome
1106 of the soybean aphid, *Aphis glycines*. *Insect Biochemistry and Molecular Biology*, *123*, 103405.
1107 https://doi.org/10.1016/j.ibmb.2020.103405
1108 Liu, S., Sappington, T. W., Coates, B. S., & Bonning, B. C. (2021). Nudivirus Sequences Identified from
1109 the Southern and Western Corn Rootworms (Coleoptera: Chrysomelidae). *Viruses*, *13*(2),

1110 Article 2. https://doi.org/10.3390/v13020269
1111 Loni, A., Samartsev, K. G., Scaramozzino, P. L., Belokobylskij, S. A., & Lucchi, A. (2016). Braconinae
1112 parasitoids (Hymenoptera, Braconidae) emerged from larvae of Lobesia botrana (Denis &
1113 Schiffermüller) (Lepidoptera, Tortricidae) feeding on Daphne gnidium L. *ZooKeys*, *587*, 125–
1114 150. https://doi.org/10.3897/zookeys.587.8478
1115 Lorenzi, A., Ravallec, M., Eychenne, M., Jouan, V., Robin, S., Darboux, I., Legeai, F., Gosselin-Grenet, A.-
1116 S., Sicard, M., Stoltz, D., & Volkoff, A.-N. (2019). RNA interference identifies domesticated viral
1117 genes involved in assembly and trafficking of virus-derived particles in ichneumonid wasps.
1118 *PLOS Pathogens*, *15*(12), Article 12. https://doi.org/10.1371/journal.ppat.1008210
1119 Lucchi, A., Ricciardi, R., Loni, A., Cosci, F., Alvarez, A. R., Beeche, M., & Scaramozzino, P. L. (2018).
1120 Rearing Campoplex capitator Aubert in Italy and in Chile: Preliminary achievements.
1121 *IOBC/Wprs Bulletin*, *139*, 96–99.
1122 Mangeney, M., Renard, M., Schlecht-Louf, G., Bouallaga, I., Heidmann, O., Letzelter, C., Richaud, A.,
1123 Ducos, B., & Heidmann, T. (2007). Placental syncytins: Genetic disjunction between the
1124 fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proceedings of the*
1125 *National Academy of Sciences*, *104*(51), Article 51. https://doi.org/10.1073/pnas.0707873105
1126 Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel
1127 and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring
1128 of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution*, *38*(10), 4647–
1129 4654. https://doi.org/10.1093/molbev/msab199
1130 Mao, M., Simmonds, T. J., Stouthamer, C. M., Kehoe, T. M., Geib, S. M., & Burke, G. R. (2023). A
1131 chromosome scale assembly of the parasitoid wasp *Venturia canescens* provides insight into
1132 the process of virus domestication. *G3: Genes, Genomes, Genetics*, *13*(10), jkad137.
1133 https://doi.org/10.1093/g3journal/jkad137
1134 Maori, E., Tanne, E., & Sela, I. (2007). Reciprocal sequence exchange between non-retro viruses and
1135 hosts leading to the appearance of new host phenotypes. *Virology*, *362*(2), Article 2.
1136 https://doi.org/10.1016/j.virol.2006.11.038
1137 Mathé-Hubert, H., Colinet, D., Deleury, E., Belghazi, M., Ravallec, M., Poulain, J., Dossat, C., Poirié, M.,
1138 & Gatti, J.-L. (2016). Comparative venomics of Psyttalia lounsburyi and P. concolor, two olive
1139 fruit fly parasitoids: A hypothetical role for a GH1 β-glucosidase. *Scientific Reports*, *6*(1), 35873.
1140 https://doi.org/10.1038/srep35873
1141 McLachlin, J. R., & Miller, L. K. (1994). Identification and characterization of vlf-1, a baculovirus gene
1142 involved in very late gene expression. *Journal of Virology*, *68*(12), Article 12.
1143 https://doi.org/10.1128/jvi.68.12.7746-7756.1994
1144 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., &

1145 Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference
1146 in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), Article 5.
1147 https://doi.org/10.1093/molbev/msaa015
1148 Moreau, J., Villemant, C., Benrey, B., & Thiéry, D. (2010). Species diversity of larval parasitoids of the
1149 European grapevine moth (Lobesia botrana, Lepidoptera: Tortricidae): The influence of region
1150 and cultivar. *Biological Control*, *54*(3), Article 3.
1151 https://doi.org/10.1016/j.biocontrol.2010.05.019
1152 Moreau, S., & Asgari, S. (2015). Venom Proteins from Parasitoid Wasps and Their Biological Functions.
1153 *Toxins*, *7*(7), Article 7. https://doi.org/10.3390/toxins7072385
1154 Murphy, N., Banks, J. C., Whitfield, J. B., & Austin, A. D. (2008). Phylogeny of the parasitic microgastroid
1155 subfamilies (Hymenoptera: Braconidae) based on sequence data from seven genes, with an
1156 improved time estimate of the origin of the lineage. *Molecular Phylogenetics and Evolution*,
1157 *47*(1), Article 1. https://doi.org/10.1016/j.ympev.2008.01.022
1158 Papura, D., Rusch, A., Roux, P., Delbac, L., & Thiéry, D. (2016). Early detection and identification of
1159 larval parasitoids in *Lobesia botrana* using PCR-RFLP method. *Biological Control*, *103*, 95–100.
1160 https://doi.org/10.1016/j.biocontrol.2016.08.005
1161 Passarelli, A., & Guarino, L. (2007). Baculovirus Late and Very Late Gene Regulation. *Current Drug*
1162 *Targets*, *8*(10), Article 10. https://doi.org/10.2174/138945007782151324
1163 Perkin, L. C., Friesen, K. S., Flinn, P. W., & Oppert, B. (2015). Venom gland components of the
1164 ectoparasitoid wasp, Anisopteromalus calandrae. *Journal of Venom Research*, *6*, 19–37.
1165 Petersen, J. M., Bézier, A., Drezen, J.-M., & van Oers, M. M. (2022). The naked truth: An updated review
1166 on nudiviruses and their relationship to bracoviruses and baculoviruses. *Journal of*
1167 *Invertebrate Pathology*, *189*, 107718. https://doi.org/10.1016/j.jip.2022.107718
1168 Pichon, A., Bézier, A., Urbach, S., Aury, J.-M., Jouan, V., Ravallec, M., Guy, J., Cousserans, F., Thézé, J.,
1169 Gauthier, J., Demettre, E., Schmieder, S., Wurmser, F., Sibut, V., Poirié, M., Colinet, D., da Silva,
1170 C., Couloux, A., Barbe, V., … Volkoff, A.-N. (2015). Recurrent DNA virus domestication leading
1171 to different parasite virulence strategies. *Science Advances*, *1*(10), Article 10.
1172 https://doi.org/10.1126/sciadv.1501150
1173 Poirié, M., Carton, Y., & Dubuffet, A. (2009). Virulence strategies in parasitoid Hymenoptera as an
1174 example of adaptive diversity. *Comptes Rendus Biologies*, *332*(2–3), Article 2–3.
1175 https://doi.org/10.1016/j.crvi.2008.09.004
1176 Poirié, M., Colinet, D., & Gatti, J.-L. (2014). Insights into function and evolution of parasitoid wasp
1177 venoms. *Current Opinion in Insect Science*, *6*, 52–60.
1178 https://doi.org/10.1016/j.cois.2014.10.004
1179 Quicke, D. L. J., & Butcher, B. A. (2021). Review of Venoms of Non-Polydnavirus Carrying Ichneumonoid

1180 Wasps. *Biology*, *10*(1), Article 1. https://doi.org/10.3390/biology10010050
1181 Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features.
1182 *Bioinformatics*, *26*(6), Article 6. https://doi.org/10.1093/bioinformatics/btq033
1183 Ratnasingham, S., & Hebert, P. D. N. (2007). BARCODING: Bold: The Barcode of Life Data System
1184 (http://www.barcodinglife.org): BARCODING. *Molecular Ecology Notes*, *7*(3), Article 3.
1185 https://doi.org/10.1111/j.1471-8286.2007.01678.x
1186 Reineke, A., Asgari, S., Ma, G., Beck, M., & Schmidt, O. (2002). Sequence analysis and expression of a
1187 virus-like particle protein, VLP2, from the parasitic wasp *Venturia canescens*. *Insect Molecular*
1188 *Biology*, *11*(3), Article 3.
1189 Reineke, A., Asgari, S., & Schmidt, O. (2006). Evolutionary origin of *Venturia canescens* virus-like
1190 particles. *Archives of Insect Biochemistry and Physiology*, *61*(3), Article 3.
1191 https://doi.org/10.1002/arch.20113
1192 Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential
1193 expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), Article 1.
1194 https://doi.org/10.1093/bioinformatics/btp616
1195 Rohrmann, G. F. (2014). Baculovirus nucleocapsid aggregation (MNPV vs SNPV): An evolutionary
1196 strategy, or a product of replication conditions? *Virus Genes*, *49*(3), Article 3.
1197 https://doi.org/10.1007/s11262-014-1113-5
1198 Rotheram, S. (1967). Immune Surface of Eggs of a Parasitic Insect. *Nature*, *214*(5089), Article 5089.
1199 https://doi.org/10.1038/214700a0
1200 Rotheram, S. (1973). The surface of the egg of a parasitic insect. I. The surface of the egg and first instar
1201 larva of *Nemeritis*. *Proceedings of the Royal Society of London. Series B. Biological Sciences*,
1202 *183*(1071), Article 1071. https://doi.org/10.1098/rspb.1973.0012
1203 Salt, G. (1965). Experimental studies in insect parasitism XIII. The haemocytic reaction of a caterpillar
1204 to eggs of its habitual parasite. *Proceedings of the Royal Society of London. Series B. Biological*
1205 *Sciences*, *162*(988), Article 988. https://doi.org/10.1098/rspb.1965.0040
1206 Santos, B. F., Klopfstein, S., Whitfield, J. B., & Sharanowski, B. J. (2022). Many evolutionary roads led
1207 to virus domestication in ichneumonoid parasitoid wasps. *Current Opinion in Insect Science*,
1208 *50*, 100861. https://doi.org/10.1016/j.cois.2021.12.001
1209 Scaramozzino, P. L., Di Giovanni, F., Loni, A., Ricciardi, R., & Lucchi, A. (2018). Updated list of the insect
1210 parasitoids (Insecta, Hymenoptera) associated with Lobesia botrana (Denis &
1211 Schiffermüller, 1775) (Lepidoptera, Tortricidae) in Italy. 2. Hymenoptera, Ichneumonidae,
1212 Anomaloninae and Campopleginae. *ZooKeys*, *772*, 47–95.
1213 https://doi.org/10.3897/zookeys.772.25288
1214 Serbielle, C., Dupas, S., Perdereau, E., Héricourt, F., Dupuy, C., Huguet, E., & Drezen, J.-M. (2012).

Evolutionary mechanisms driving the evolution of a large polydnavirus gene family coding for protein tyrosine phosphatases. *BMC Evolutionary Biology*, *12*(1), 253. https://doi.org/10.1186/1471-2148-12-253

Sharanowski, B. J., Ridenbaugh, R. D., Piekarski, P. K., Broad, G. R., Burke, G. R., Deans, A. R., Lemmon, A. R., Moriarty Lemmon, E. C., Diehl, G. J., Whitfield, J. B., & Hines, H. M. (2021). Phylogenomics of Ichneumonoidea (Hymenoptera) and implications for evolution of mode of parasitism and viral endogenization. *Molecular Phylogenetics and Evolution*, *156*, 107023. https://doi.org/10.1016/j.ympev.2020.107023

Singh, U., & Wurtele, E. S. (2021). orfipy: A fast and flexible tool for extracting ORFs. *Bioinformatics*, *37*(18), Article 18. https://doi.org/10.1093/bioinformatics/btab090

Stoltz, D. B. (1990). Evidence for Chromosomal Transmission of Polydnavirus DNA. *Journal of General Virology*, *71*(5), 1051–1056. https://doi.org/10.1099/0022-1317-71-5-1051

Strand, M. R., & Burke, G. R. (2012). Polydnaviruses as Symbionts and Gene Delivery Systems. *PLoS Pathogens*, *8*(7), Article 7. https://doi.org/10.1371/journal.ppat.1002757

Strand, M. R., & Burke, G. R. (2014). Polydnaviruses: Nature's Genetic Engineers. *Annual Review of Virology*, *1*(1), Article 1. https://doi.org/10.1146/annurev-virology-031413-085451

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, *34*(Web Server), Article Web Server. https://doi.org/10.1093/nar/gkl315

Tarlinton, R. E., Meers, J., & Young, P. R. (2006). Retroviral invasion of the koala genome. *Nature*, *442*(7098), Article 7098. https://doi.org/10.1038/nature04841

Teng, Z.-W., Xiong, S.-J., Xu, G., Gan, S.-Y., Chen, X., Stanley, D., Yan, Z.-C., Ye, G.-Y., & Fang, Q. (2017). Protein Discovery: Combined Transcriptomic and Proteomic Analyses of Venom from the Endoparasitoid Cotesia chilonis (Hymenoptera: Braconidae). *Toxins*, *9*(4), 135. https://doi.org/10.3390/toxins9040135

Theopold, U., Krause, E., & Schmidt, O. (1994). Cloning of a VLP-protein coding gene from a Parasitoid Wasp *Venturia canescens*. *Archives of Insect Biochemistry and Physiology*, *26*(2–3), Article 2–3. https://doi.org/10.1002/arch.940260207

Thézé, J., Bézier, A., Periquet, G., Drezen, J.-M., & Herniou, E. A. (2011). Paleozoic origin of insect large dsDNA viruses. *Proceedings of the National Academy of Sciences*, *108*(38), Article 38. https://doi.org/10.1073/pnas.1105580108

Thiéry, D., Delbac, L., Villemant, C., & Moreau, J. (2011). Control of grape berry moth larvae using parasitoids: Should it be developed? *IOBC-WPRS Bulletin*, *67*(26), Article 26. https://doi.org/10.13140/2.1.3313.7442

Vanarsdall, A. L., Okano, K., & Rohrmann, G. F. (2006). Characterization of the Role of Very Late

1250 Expression Factor 1 in Baculovirus Capsid Structure and DNA Processing. *Journal of Virology*,
1251 *80*(4), Article 4. https://doi.org/10.1128/JVI.80.4.1724-1733.2006
1252 Volkoff, A.-N., & Cusson, M. (2020). The Unconventional Viruses of Ichneumonid Parasitoid Wasps.
1253 *Viruses*, *12*(10), Article 10. https://doi.org/10.3390/v12101170
1254 Volkoff, A.-N., Drezen, J.-M., Cusson, M., & Webb, B. A. (2012). The Organization of Genes Encoding
1255 Ichnovirus Structural Proteins. In *Parasitoid Viruses* (pp. 33–45). Elsevier.
1256 https://doi.org/10.1016/B978-0-12-384858-1.00003-5
1257 Volkoff, A.-N., Jouan, V., Urbach, S., Samain, S., Bergoin, M., Wincker, P., Demettre, E., Cousserans, F.,
1258 Provost, B., Coulibaly, F., Legeai, F., Béliveau, C., Cusson, M., Gyapay, G., & Drezen, J.-M.
1259 (2010). Analysis of Virion Structural Components Reveals Vestiges of the Ancestral Ichnovirus
1260 Genome. *PLoS Pathogens*, *6*(5), Article 5. https://doi.org/10.1371/journal.ppat.1000923
1261 Wang, L., Fang, Q., Qian, C., Wang, F., Yu, X.-Q., & Ye, G. (2013). Inhibition of host cell encapsulation
1262 through inhibiting immune gene expression by the parasitic wasp venom calreticulin. *Insect*
1263 *Biochemistry and Molecular Biology*, *43*(10), 936–946.
1264 https://doi.org/10.1016/j.ibmb.2013.07.010
1265 Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C. W., Jones, B. R., Bradley,
1266 T., Zhu, H., Guan, Y., Jiang, Y., & Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input
1267 and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution*,
1268 *37*(2), 599–603. https://doi.org/10.1093/molbev/msz240
1269 Wang, Y., Kleespies, R. G., Huger, A. M., & Jehle, J. A. (2007). The Genome of *Gryllus bimaculatus*
1270 Nudivirus Indicates an Ancient Diversification of Baculovirus-Related Nonoccluded Nudiviruses
1271 of Insects. *Journal of Virology*, *81*(10), Article 10. https://doi.org/10.1128/JVI.02781-06
1272 Wang, Y., R.P., O., & A., J. (2012). Nudivirus Genomics and Phylogeny. In M. Garcia (Ed.), *Viral*
1273 *Genomes—Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus*
1274 *Interactions* (pp. 33–52). InTech. https://doi.org/10.5772/27793
1275 Wey, B., Heavner, M. E., Wittmeyer, K. T., Briese, T., Hopper, K. R., & Govind, S. (2020). Immune
1276 Suppressive Extracellular Vesicle Proteins of *Leptopilina heterotoma* Are Encoded in the Wasp
1277 Genome. *G3 Genes|Genomes|Genetics*, *10*(1), 1–12. https://doi.org/10.1534/g3.119.400349
1278 Wong, E. S. W., & Belov, K. (2012). Venom evolution through gene duplications. *Gene*, *496*(1), Article
1279 1. https://doi.org/10.1016/j.gene.2012.01.009
1280 Wu, W., Lin, T., Pan, L., Yu, M., Li, Z., Pang, Y., & Yang, K. (2006). *Autographa californica* Multiple
1281 Nucleopolyhedrovirus Nucleocapsid Assembly Is Interrupted upon Deletion of the *38K* Gene.
1282 *Journal of Virology*, *80*(23), Article 23. https://doi.org/10.1128/JVI.01155-06
1283 Yang, S., & Miller, L. K. (1998). Expression and Mutational Analysis of the Baculovirus Very Late Factor
1284 1 (vlf-1) Gene. *Virology*, *245*(1), Article 1. https://doi.org/10.1006/viro.1998.9152

1285 Yang, S., & Miller, L. K. (1999). Activation of Baculovirus Very Late Promoters by Interaction with Very
1286 Late Factor 1. *Journal of Virology*, *73*(4), Article 4. https://doi.org/10.1128/JVI.73.4.3404-
1287 3409.1999
1288 Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An r package for visualization and
1289 annotation of phylogenetic trees with their covariates and other associated data. *Methods in*
1290 *Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628
1291 Zhang, G., Schmidt, O., & Asgari, S. (2006). A calreticulin-like protein from endoparasitoid venom fluid
1292 is involved in host hemocyte inactivation. *Developmental & Comparative Immunology*, *30*(9),
1293 756–764. https://doi.org/10.1016/j.dci.2005.11.001
1294 Zhang, J.-H., Washburn, J. O., Jarvis, D. L., & Volkman, L. E. (2004). *Autographa californica* M
1295 nucleopolyhedrovirus early GP64 synthesis mitigates developmental resistance in orally
1296 infected noctuid hosts. *Journal of General Virology*, *85*(4), Article 4.
1297 https://doi.org/10.1099/vir.0.19773-0
1298 Zhang, Y., Wang, J., & Han, G.-Z. (2020). Chalcid wasp paleoviruses bridge the evolutionary gap
1299 between bracoviruses and nudiviruses. *Virology*, *542*, 34–39.
1300 https://doi.org/10.1016/j.virol.2020.01.007
1301 Zheng, Q., Shen, Y., Kon, X., Zhang, J., Feng, M., & Wu, X. (2017). Protein–protein interactions of the
1302 baculovirus per os infectivity factors (PIFs) in the PIF complex. *Journal of General Virology*,
1303 *98*(4), Article 4. https://doi.org/10.1099/jgv.0.000730

Table 1: Summary table of CcapiENV gene information. Putative functions were predicted from gene functions described in baculoviruses. Functional nudivirus core genes are indicated in bold and pseudogenized nudivirus core genes are indicated with an asterisk. Names between parentheses correspond to alternative names of the genes found in the literature. TPMs and PSM correspond to transcripts per million and peptide-spectrum matches respectively. A complete version of this table is available in supplementary data (Table S3).

Function in baculoviruses	Gene name	Contig	Nudiviral cluster	Position in the contig	Gene length in bp	Mean of normalized TPMs in ovaries	Presence in VLPs	Peptide abundance (PSM)	Protein size in aa
DNA amplification	***helicase***	contig 13	2	complement(6986111 - 6989746)	3636	60.10636905			1212
Viral transcription	***p47***	contig 2	1	complement(15493852 - 15494922)	1071	202.0627154			357
	lef-4	contig 13	2	complement(6977003 - 6978238)	1236	307.9753009			412
	lef-8	contig 13	2	6982207 - 6985356	3150	190.7332664			1050
	lef-5	contig 328	4	complement(708144 - 708470)	327	390.8433835			109
	lef-9	contig 328	4	complement(727668 - 729752)	2085	380.7544624			695
Envelope components and infectiosity	***pif-4 (19 Kda)***	contig 328	4	706840 - 707595	756	10234.6896	X	39	252
	p33 (ac92)	contig 2	1	15496092 - 15497036	945	4562.742796	X	49	315
	pif-2	contig 13	2	complement(6971652 - 6972776)	1125	1950.200664	X	51	375
	pif-6 (ac68)	contig 328	4	complement(690595 - 691005)	411	4524.153106	X	13	137
	p74 (pif-0)	contig 328	4	695558 - 697498	1941	1979.354037	X	23	647
	vp91 (pif-8)	contig 328	4	complement(713819 - 715891)	2073	1898.483467	X	71	691
	pif-5-1 (odv-e56)	contig 328	4	725333 - 726583	1251	8407.359441	X	104	417
	pif-5-2 (odv-e56)	contig 328	4	733696 - 734952	1257	2238.100222	X	127	419
	pif-3	contig 328	4	744312 - 744905	594	5389.304417	X	33	198
	pif-1	contig 13	2	6978638 - 6980113	1476	1358.325472	X	24	492
	Ac81	contig 328	4	701318 - 701800	483	870.8774304			161
Morphogenesis and particle assembly	***vlf-1***	contig 328	4	709361 - 710638	1278	785.8570436			426
Unknown	*OrNVorf59-like*	contig 328	4	707602 - 708123	522	9224.822369	X	43	174

OrNVorf54-like	contig 2	1	complement(15472150 - 15473337)	1188	3589.443084	X	49	396
OrNVorf41-like-1 (11K)	contig 2	1	15475263 - 15475550	288	32236.6679	X	13	96
OrNVorf41-like-2 (11K)	contig 2	1	15476320 - 15476727	408	20804.23563	X	33	136
OrNVorf41-like-3 (11K)	contig 2	1	15479909 - 15480358	450	41424.23901			150
OrNVorf41-like-4 (11K)	contig 2	1	15481182 - 15481586	405	55874.69226	X	13	135
OrNVorf44-like	contig 2	1	15483923 - 15484489	567	11612.46173	X	66	189
OrNVorf46-like	contig 2	1	complement(15486115 - 15487905)	1791	151.0870697			597
OrNVorf136-like	contig 2	1	complement(15488716 - 15489168)	453	3212.888481			151
OrNVorf23-like	contig 2	1	15489783 - 15490574	792	261.6013813	X	47	264
OrNVorf19-like	contig 2	1	complement(15494837 - 15495235)	399	420.6676754			133
OrNVorf47-like	contig 13	2	6975948 - 6976769	822	2839.655997	X	18	274
OrNVorf73-like	contig 13	2	6992759 - 6993544	786	242.6407675			262
OrNVorf41-like-5 (11K)	contig 326	3	complement(2612826 - 2613125)	300	11433.84229			100
OrNVorf41-like-6 (11K)	contig 326	3	complement(2613850 - 2614170)	321	12768.01194			107
OrNVorf138-like	contig 328	4	complement(686054 - 686347)	294	2268.386529			98
OrNVorf61-like	contig 328	4	complement(689155 - 689526)	372	10869.81904	X	25	124
OrNVorf76-like	contig 328	4	691432 - 691659	228	978.3012378			76
OrNVorf25-like	contig 328	4	698675 - 700330	1656	3713.565782	X	191	552
OrNVorf27-like	contig 328	4	710904 - 712094	1191	2834.742838	X	20	397
OrNVorf90-like	contig 328	4	complement(717857 - 721093)	3237	225.2549736	X	35	1079
OrNVorf63-like	contig 328	4	complement(723007 - 723255)	249	2050.01484			83
OrNVorf79-like	contig 328	4	731525 - 731737	213	1646.832857	X	11	71
OrNVorf118-like	contig 328	4	736112 - 736948	837	3367.630323	X	97	279
OrNVorf119-like	contig 328	4	736958 - 738058	1101	7250.849857	X	80	367
OrNVorf120-like	contig 328	4	complement(738811 - 740787)	1977	1880.881495	X	120	659
OrNVorf122-like	contig 328	4	741734 - 742414	681	364.2703585			227
OrNVorf123-like	contig 328	4	743586 - 743909	324	6143.331885	X	17	108
OrNVorf18-like	contig 328	4	746646 - 748046	1401	6663.775966	X	109	467

Pseudogenes	pseudo vp39 *	contig 2	1	15465017 - 15465368
	pseudo OrNVorf99-like	contig 4	NA	6410523 - 6410915
	pseudo OrNVorf117-like	contig 9	NA	1382035 - 1382601
	pseudo OrNVorf139-like	contig 9	NA	6965367- 6965623
	pseudo FEN-1 *	contig 12	NA	726667 - 726882, 727180 - 727562
	pseudo DNApol *	contig 13	NA	507908 - 507956, 508176 - 508868
	pseudo OrNVorf99-like	contig 13	NA	5114850 - 5115368
	pseudo pif-4 *	contig 13	2	6989881 - 6990024
	pseudo pif-1 *	contig 13	2	6994508 - 6994810
	pseudo helicase *	contig 13	2	complement(6995640 - 6995816)
	pseudo OrNVorf62-like	contig 13	NA	7033194 - 7033540
	pseudo 38K *	contig 94	NA	29063 - 29227
	pseudo integrase *	contig 322	NA	7794008 - 7794318, 7794375 - 7794869
	pseudo OrNVorf22-like (p6.9) *	contig 322	NA	12698893 - 12699488
	pseudo OrNVorf128-like	contig 322	NA	14126845 - 14127067
	pseudo OrNVorf27-like	contig 326	3	2618461 - 2618893
	pseudo lef-8 *	contig 504	NA	3017 - 3559
	pseudo 38K *	contig 508	NA	23980 - 24171

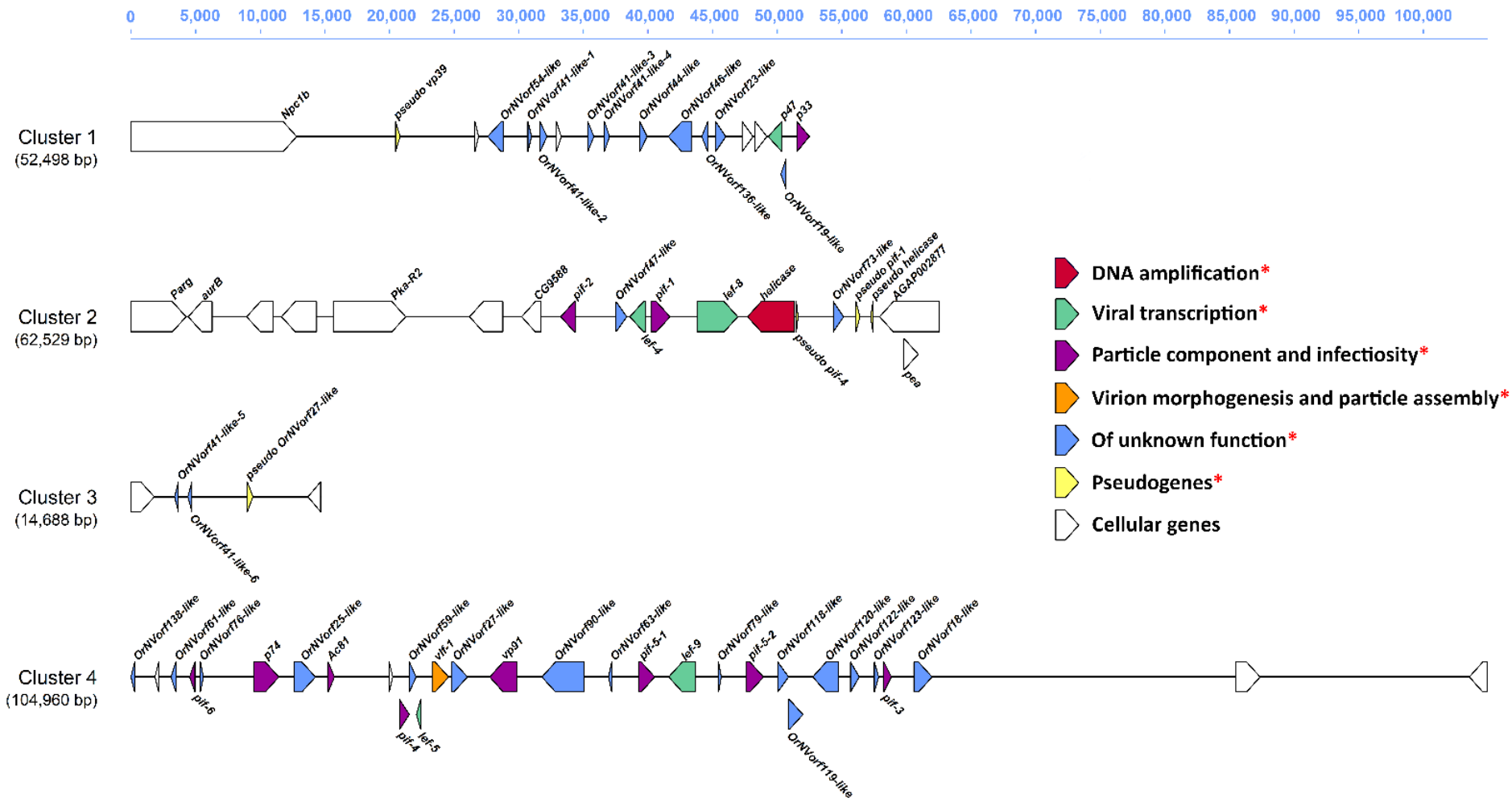


Figure 1: The nudiviral clusters of CcapiENV. Sequences of nudiviral origin are indicated by a red asterisk in the legend. Nudiviral genes were colourised according to their function described in baculoviruses. Cellular genes correspond to genes of non-nudiviral origin (with no nudiviral hit). The scale represents the cluster length in base pairs

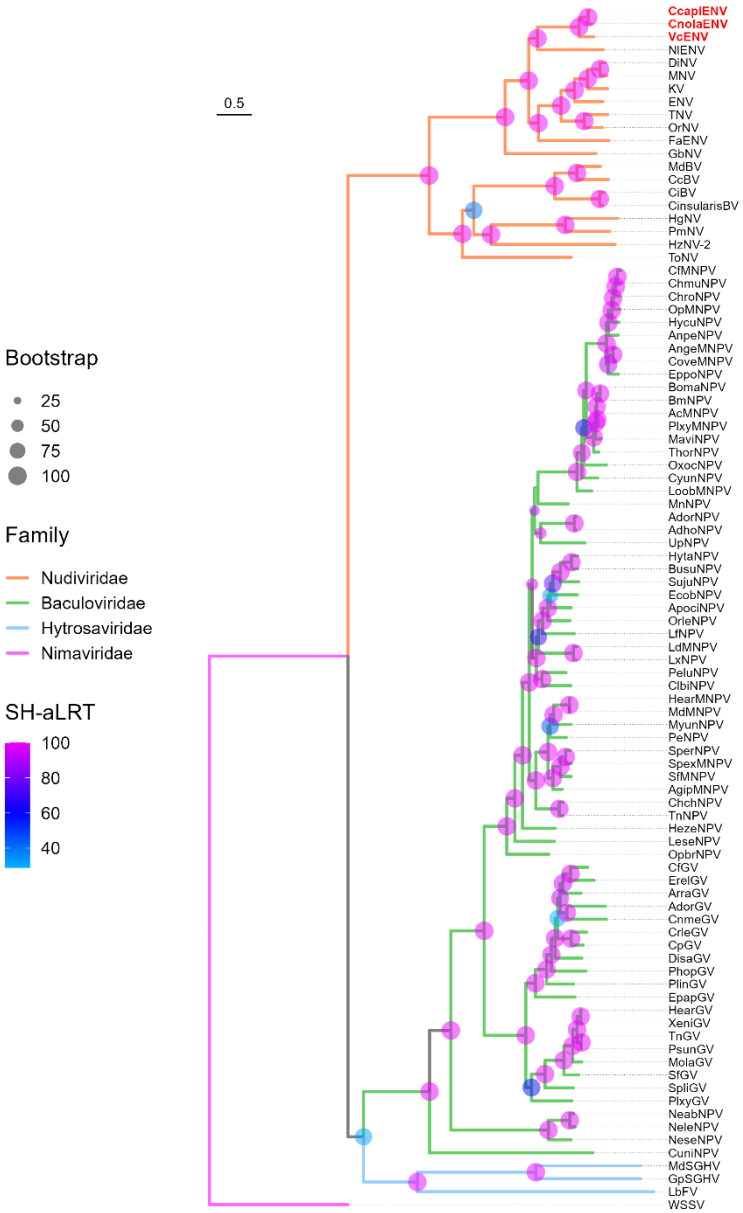


Figure 2: Phylogenetic tree of Naldaviricetes viruses. The tree includes 20 nudiviruses, 69 baculoviruses, 3 hytrosaviruses (including LbFV for Leptopilina boulardi filamentous virus) and a nimavirus (WSSV for White spot syndrome virus) set as the outgroup. In total, 34 genes (21 baculovirus core genes and 13 alphanudivirus core genes) were used for the tree construction, using the maximum likelihood criterion. Branch supports are given by the SH-aLRT (Shimodaira-Hasegawa approximate likelihood ratio test), corresponding to the colour of the node circles and bootstrap values, corresponding to their size. The scale bar represents the phylogenetic distance (percentage of genetic variation). Campoplex capitator Endogenous Nudivirus and Campoplex nolae Endogenous Nudivirus are called respectively CcapiENV and CnolaENV. Nudiviruses endogenized in Campopleginae wasps are mentioned in red. Virus and sequence information is given in supplementary data (Table S4).

Table 2: Nudivirus genes in the *C. capitator, C. nolae* and *V. canescens* genomes. Genes found in *C. capitator* (in orange) and *C. nolae* (in blue) were compared with those of VcENV (in green). Genes and pseudogenes are indicated with solid dots and crossed dots respectively. Copies are indicated by multiple dots. "na" indicates that no copy of the corresponding gene was found. Alternative gene names are given between parentheses.

Function in baculoviruses	Gene name	CcapiENV	CnolaENV	VcENV
DNA amplification	*DNApol*	⊗	⊗	⊗
	FEN-1	⊗	⊗	⊗
	helicase	●⊗	●	●
	helicase2	na	na	na
	integrase	⊗	⊗	⊗
Viral transcription	*lef-4*	●	●	●⊗
	lef-5	●	●	●
	lef-8	●⊗	●	●
	lef-9	●	●	●
	p47	●	●	●
Envelope components, infectiosity	*ac81*	●	●	●●●
	p33 (ac92)	●	●	●
	p74 (pif-0)	●	●	●
	pif-1	●⊗	●	●⊗
	pif-2	●	●	●
	pif-3	●	●	●
	pif-4 (19 Kda)	●⊗	●⊗	●
	pif-5 (odv-e56)	●●	●●	●●●●
	pif-6 (ac68)	●	●	●
	vp91(pif-8)	●	●	●
Morphogenesis, particle assembly	*38K*	⊗	na	⊗
	vlf-1	●	●	⊗
	vp39	⊗	⊗	⊗
	OrNVorf22-like (p6.9)	⊗	⊗	⊗
Nucleotid metabolism	*tk1*	na	na	na
	tk2	na	na	na
	tk3	na	na	na
Unknown	*OrNVorf118-like*	●	●	●
	OrNVorf119-like	●	●	●
	OrNVorf120-like	●	●	●
	OrNVorf122-like	●	●	●
	OrNVorf123-like	●	●	●
	OrNVorf136-like	●	●	●
	OrNVorf138-like	●	●	●

OrNVorf18-like	●	●	●
OrNVorf19-like	●	●	●
OrNVorf23-like	●	●	●
OrNVorf25-like	●	●	●
OrNVorf27-like	●⊗	●⊗	●
OrNVorf41-like (11K)	●●●●●●	●●●●●●	●●●●●
OrNVorf44-like	●	●	●
OrNVorf46-like	●	●	●
OrNVorf47-like	●	●	●●●
OrNVorf54-like	●	●	●
OrNVorf59-like	●	●⊗	●
OrNVorf63-like	●	●	●
OrNVorf73-like	●	●	●
OrNVorf76-like	●	●	●
OrNVorf79-like	●	●	●
OrNVorf90-like	●	●	●
OrNVorf2-like	na	na	⊗
OrNVorf9-like	na	⊗	⊗
OrNVorf62-like	⊗	⊗	⊗
OrNVorf99-like	⊗	⊗	⊗
OrNVorf117-like	⊗	⊗	⊗
OrNVorf128-like	⊗	⊗	⊗
OrNVorf130-like	na	na	⊗
OrNVorf139-like	⊗	⊗	⊗

1328

1329

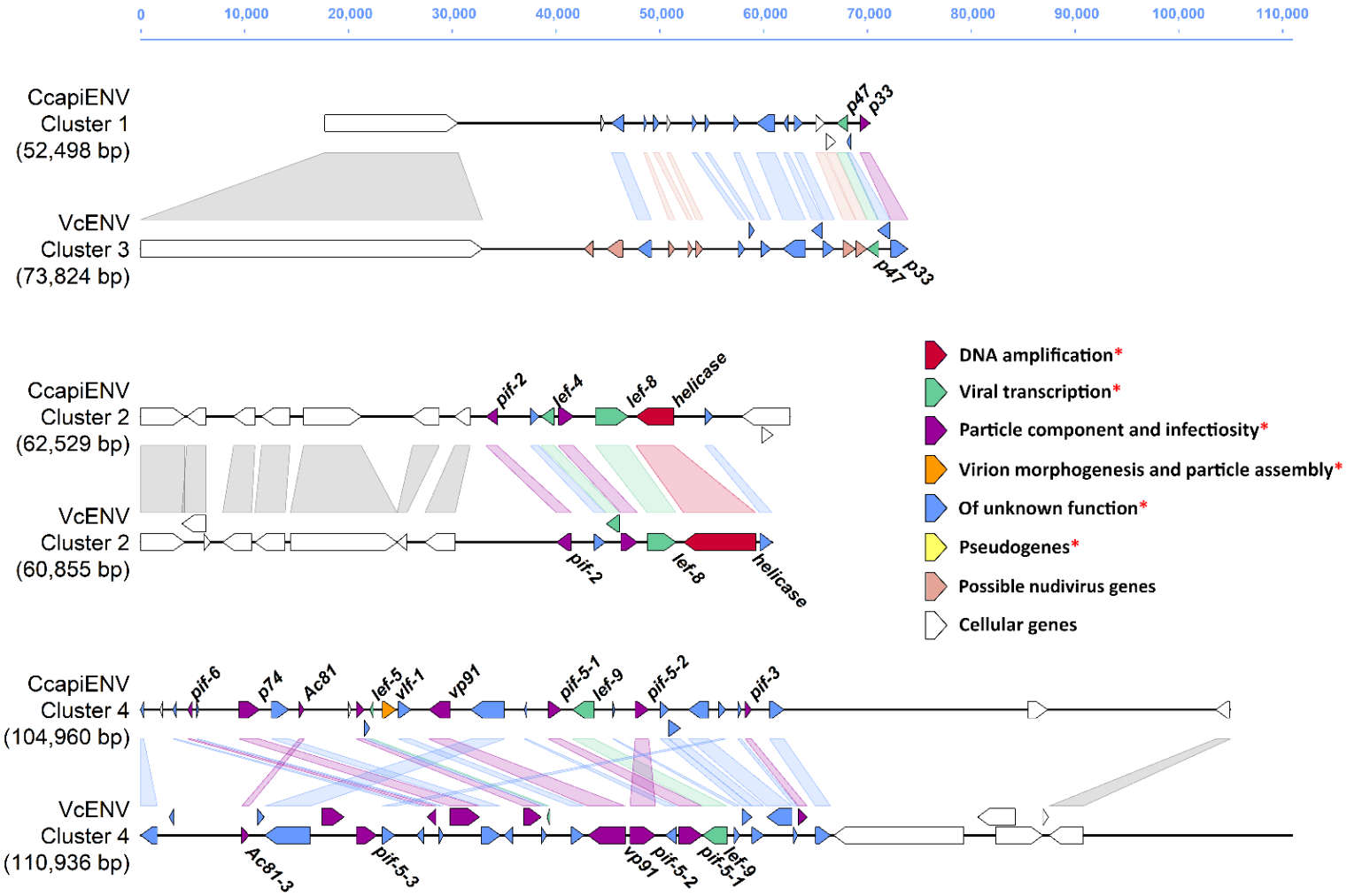
1330

1331

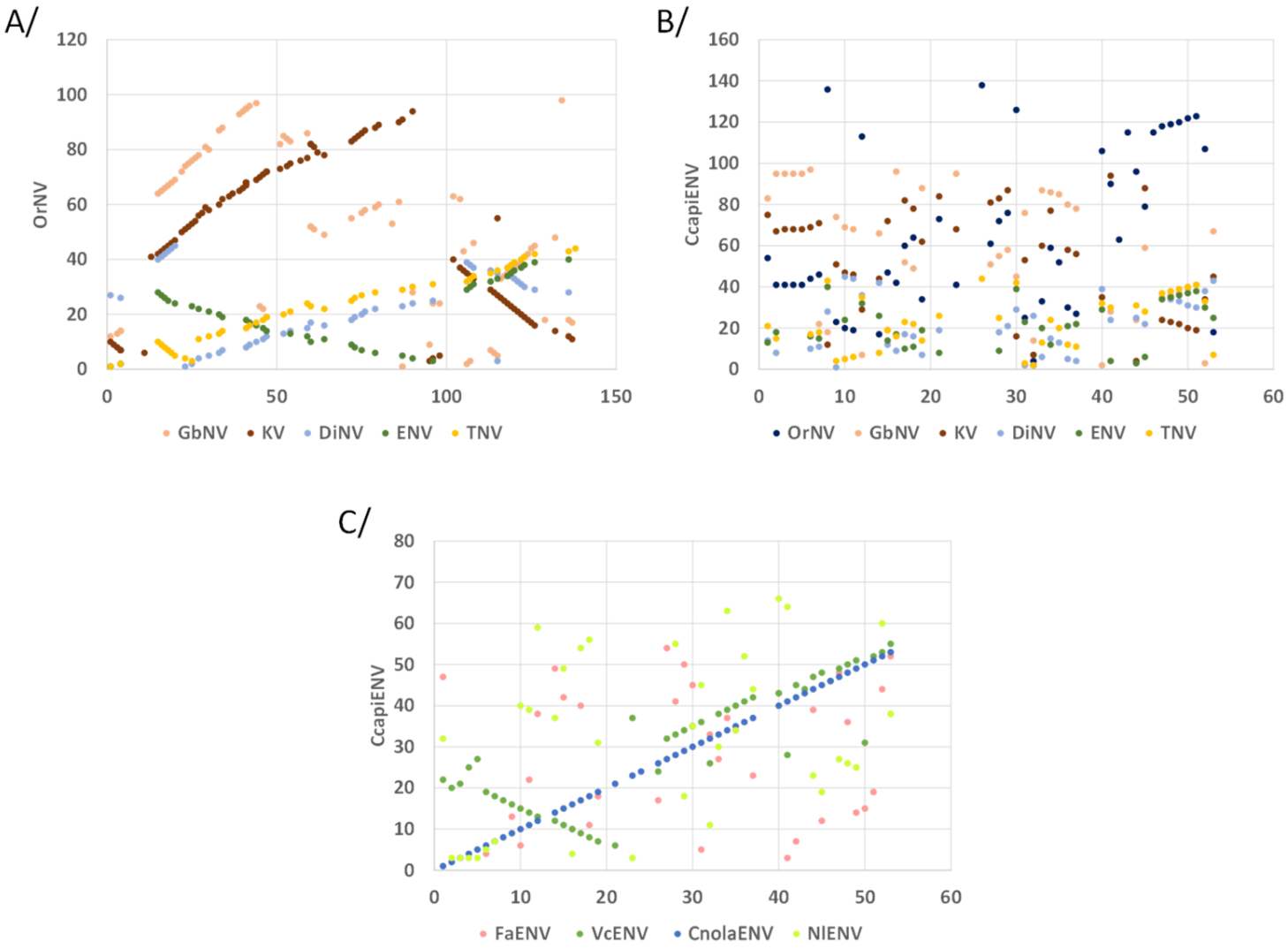
1332

1333

1334



1335 Figure 3: Syntenies between the CcapiENV clusters and VcENV clusters (from Mao et al., 2023). Models were colourised according to their
1336 function. Light pink models correspond to genes annotated in Mao and collaborators (Mao et al., 2023) which have no match in the NCBI
1337 database, but have nevertheless been annotated as potential virus genes. The scale represents the length in base pairs.



1338

1339 Figure 4: Parity plots comparing nudivirus genes order. **A** - Parity plot between the genomes
1340 of Oryctes rhinoceros nudivirus (OrNV) on the horizontal axis and other exogenous nudiviruses
1341 on the abscissa. The other nudiviruses are Gryllus bimaculatus nudivirus (GbNV), Kallithea
1342 virus (KV) Drosophila innubia nudivirus (DiNV), Esparto virus (ENV) and Tomelloso virus (TNV).
1343 **B** - Parity plot between the genomes of CcapiENV (ordinate) and the previous exogenous
1344 nudiviruses (abscissa). **C** - Parity plot between the genomes of CcapiENV and other
1345 endogenized nudiviruses (abscissa). The other endogenized nudiviruses are Fopius arisanus
1346 endogenous nudivirus (FaENV), VcENV, CnolaENV and Nilaparvata lugens endogenous
1347 nudivirus (NlENV).

1348

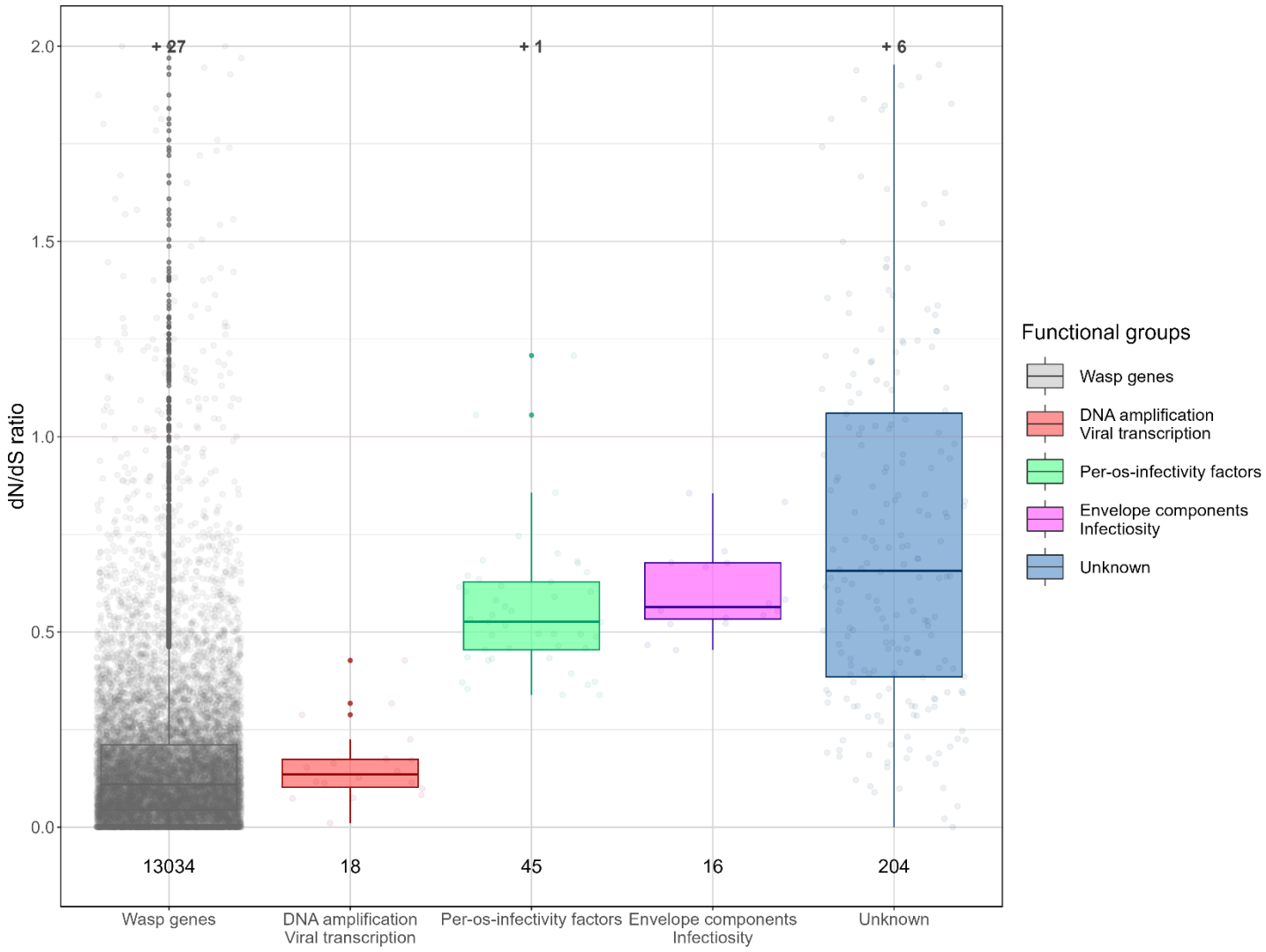


Figure 5: Estimation of selective pressures acting on wasp and endogenized nudivirus genes. dN/dS ratio values were calculated from wasp (in grey) and nudivirus genes of *C. capitator*, *C. nolae* and *V. canescens*. Numbers below each boxplot represent the number of sequences used for dN/dS ratio calculation for each category. Numbers above each boxplot represent the number of values higher than 2, which could not be plotted. In red are the dN/dS ratio values for *helicase*, *p47* and *lef* genes, in green are the values for *pif* (*per-os-infectivity factor*) genes, in pink the values for *Ac81*, *vp91* and *p33* genes, and in blue the values for *OrNVorf-like* genes.

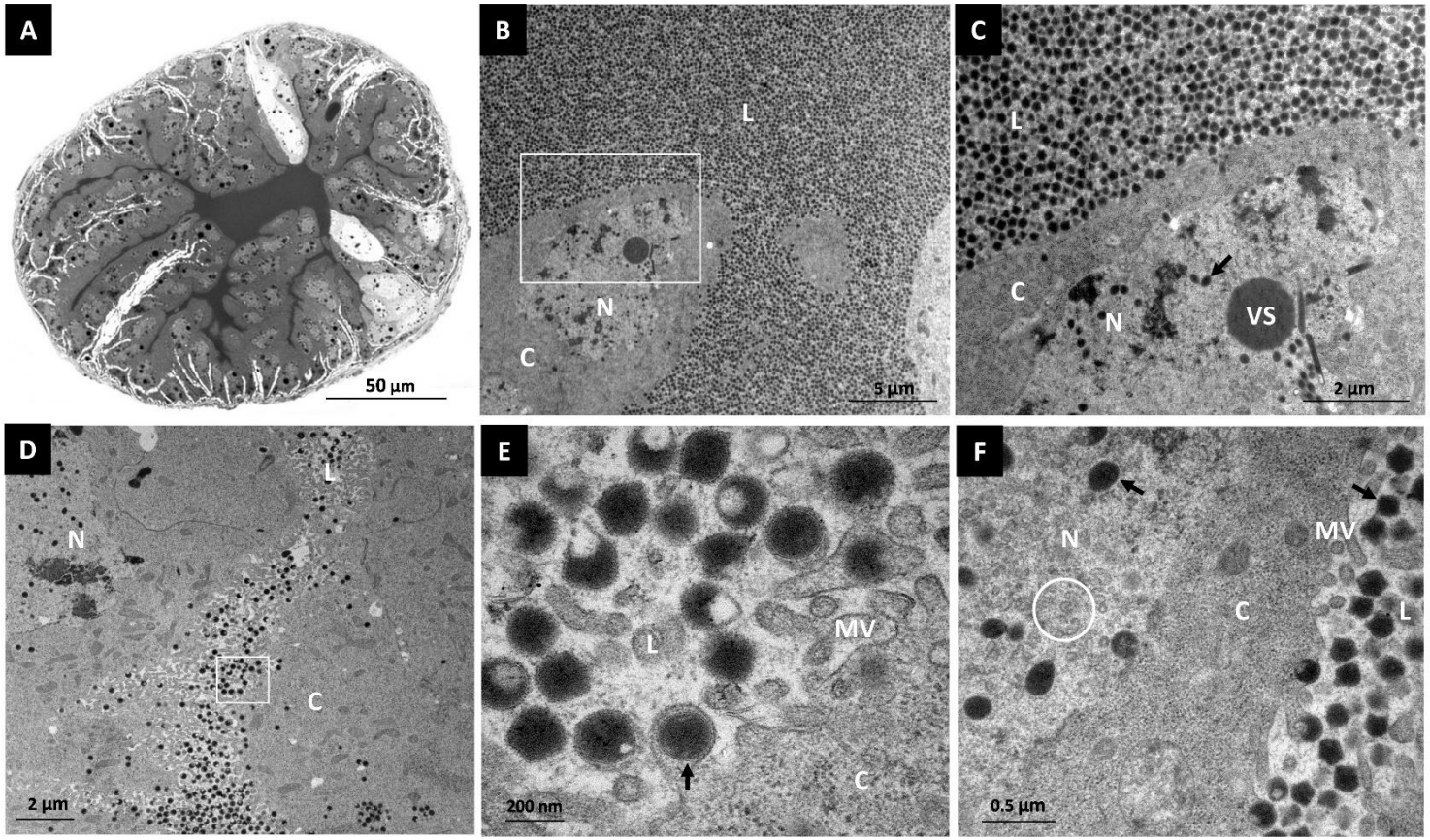


Figure 6: The calyx region of a *C. capitator* ovaries of adult wasp observed using light and electron microscopy. A - light microscopic photograph of a semi-thin section of the calyx region, stained with toluidine blue. B, D - electron microscopy photographs of the calyx region shown in A. C, E - enlarged areas (white frames) of photographs B and D, respectively. Virogenic stromas (VS) and VLPs (dark arrow) are visible in cell nuclei (N). While virogenic stromas can be seen only in the nuclei, VLPs appear to travel from the nucleus to the lumen (L) where they accumulate, passing through the cytoplasm (C) and the membrane microvilli (MV). F - electron microscopy photograph showing the nucleus-cytoplasm-lumen interface where empty VLP envelopes can be seen in the nucleus (white circle), as described in *V. canescens* (Pichon et al., 2015).